

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2003, 06:17:30 ; Search time 58 Seconds
(without alignments)
863.552 Million cell updates/sec

Title: US-09-884-566A-2
Perfect score: 2650
Sequence: 1 MEVLESQSVLQWDRKLS.....LEGNETLKVVELRRVATP 521

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
al number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	445.5	16.8	515	2 A42140	box B-binding factor
2	436	16.5	516	2 A44494	CAMP-responsive el
3	303.5	11.5	688	2 T32750	hypothetical prote
4	298	11.2	365	2 T15656	hypothetical prote
5	252	9.5	699	2 T09069	probable CAMP-resp
6	167.5	6.3	741	2 I48694	probable transcrip
7	166	6.3	742	2 A49672	transcription fact
8	163	6.2	609	2 T49282	hypothetical prote
9	162.5	6.1	341	2 S26686	CAMP response elem
10	162.5	6.1	486	2 J4028	activating transcr
11	162	6.1	313	1 TVQJUN	transforming prote
12	161.5	6.1	278	2 S36101	CAMP response elem
13	160	6.0	322	2 T08592	TGACG-motif-bindin
14	159.5	6.0	296	1 TVFVJN	transforming prote
15	158.5	6.0	229	2 A37944	CAMP response elem
16	158.5	6.0	271	2 S12560	transcription fact
17	158	6.0	344	2 J55601	CAMP response elem
18	157.5	5.9	269	2 A46490	TCR V beta CRE-mot
19	157	5.9	772	2 A55004	transcription fact
20	156.5	5.9	310	2 I50373	jun protein - chic
21	155	5.8	322	2 T12093	TGACG-motif bindin
22	154.5	5.8	373	2 S32537	erythroid transcri
23	152.5	5.8	373	2 A49671	gene NF-E2 protein
24	152	5.7	334	1 TVMSJA	transcription fact
25	151.5	5.7	489	2 S50548	hypothetical prote
26	151.5	5.7	373	2 A54692	transcription fact
27	151.5	5.7	483	2 S32741	transcription fact
28	150	5.7	327	2 S22298	CAMP response elem
29	150	5.7	341	1 TVMSJD	transforming prote

30	150	5.7	341	2 JC4051	jun-D protein - ra
31	150	5.7	589	2 I59340	transcription fact
32	150	5.7	679	2 S48437	hypothetical prote
33	149.5	5.6	325	2 S23007	CAMP response elem
34	149.5	5.6	1844	1 RRWPTM	genome polyprotein
35	149	5.6	334	2 S12742	transcription fact
36	149	5.6	448	2 A42026	CAMP response elem
37	149	5.6	456	2 B42026	cyclic AMP respons
38	148.5	5.6	326	2 T08591	TGACG-motif bindin
39	148.5	5.6	328	2 A35663	CAMP response elem
40	148	5.6	360	2 T03373	probable G-box bin
41	148	5.6	372	2 T10472	G-box binding prot
42	147.5	5.6	341	2 S20827	CAMP response elem
43	147.5	5.6	505	1 S05380	transcription fact
44	147	5.5	326	2 A40120	CAMP-responsive en
45	147	5.5	333	2 T26808	hypothetical prote

ALIGNMENTS

RESULT 1

A42140
box B-binding factor-2 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A42140; S24542
R:Abel, T.; Bhatt, R.; Maniatis, T.
Genes Dev. 6, 466-480, 1992
A:Title: A Drosophila CREB/ATF transcriptional activator binds to both fat body- and liv
A:Reference number: A42140; MUID:92192458; PMID:1532159
A:Accession: A42140
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-515 <ABR>
A:Cross-references: EMBL:X64429; NID:g11063; PIDN:CAA45771.1; PID:g11064
A:Note: sequence extracted from NCI backbone (NCBIP:88160)
C:Genetics:
A:Gene: FlyBase:Bbbf2
A:Cross-references: FlyBase:FBgn0004848
A:Superfamily: fos/jun DNA-binding domain homology
C:Keywords: DNA binding; nucleus; transcription regulation
F:432-475/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match	16.8%;	Score	445.5;	DB 2;	Length	515;	
Best Local Similarity	33.8%;	Pred. No.	8.4e-19;				
Matches	142;	Conservative	54;	Mismatches	117;	Gaps	16;
QY	26	EFEALMYH--THFSELDLDFSNVGLQLLDPFLSEKSESMVEFSPSTPAPLQAEHS	82				
Db	111	DPKCLTFHVPPTATP-----ISRLSSNPALN-----TSVADLTRSSGL	149				
QY	83	YSL-SEERTQSPFTHATSD-----SFNDEVESEKWL-----STEPS	122				
Db	150	OSLQAHQHGGSGSHVVANLEHFLQFOLYDNCSSVSLRDGSMSPDICSDEIDE	209				
QY	123	ATIKKEPITEEQGLVPS-----VTLITATSTP--FEKESPLDMNAGDSSCQ	171				
Db	210	SAIKDEPMSPDSCSPASTSQSSSQHQLSLNLAHQEMLEFPHKCGLLLTASSNSN-N	268				
QY	172	TLIPK-----IKLEPHEVDQFLNFPKASVDQLHLPTTPPS-HSSDSE	215				
Db	269	SLIKSQQRHEQILQDNLIMAKMEIKSEKQSTNSGKSHAGYGIPTTPPSLPSDDSE	328				
QY	216	GSLSPN-----PRL-----HPFSLQASHVPR-----AMPRGPSALST--	248				
Db	329	GNLSPEHLFRHCRPTQFPFLMPIQFAVSHPYGSATAASITRSSGSASGSSTGSTVT	388				
QY	249	-----SPLLTAPKLOSGGPLVLTTEEBKRTLVAEGYPIPTKPLTKSEKALKKIRK	301				
Db	389	TTRQPIHTPLISSQPK-GSTGTLLTEEBKRTLVAEGYPIPKPLTKKAEKSKLIRK	447				
QY	302	IKNKISAQESRRKKKEYMDSLEKVKESCSTENLELRKKVLENTNRTLLOOLQLOTIV	361				

Matches 94; Conservative 58; Mismatches 137; Indels 123; Gaps 18;

QY 1 MEVLSEGSQVLOWDRKLSLSEPEGETEALMHTHPSSELLDFEQNVLCQL-LSDPFLSE 59
Db 3 METVES-----QDRSV-----THSVAHSSLLHMQTQISVPTLAQVSVAGSGTGR 48

QY 60 KSESMEVEPSPTSPA---PLIAQHSYSLSEPRQTS--PETHAATSDSFNDEVESEK 113
Db 49 GSPATVLQVPSQTVQVGVLTQHP-SVIGSPQIQTVQVATIAETDSDASEVIDSHK 107

QY 114 W-----YLSTEPSATIKKEPITEEO--PPGL-----VP-SVTLTITA--- 148
Db 108 RREILSRPSYRKILNELSSDVGPIKIEEKSEEGTTPNIATMAVPTSIVQSTGQVI 167

QY 149 -----ISTPFKEESPLD-----MNAGDSSCOTLIPKIKLPEHVDQFLNFPKEA 195
Db 168 ATAOGGTIQISNPGSDGVQGLQALTMNAGAPPAGATVQYAAQADGTQOF--FVPGSQ 225

QY 196 SV---DQLHLPPTPSSHSSDSEGLSPNRLHPSLSQAHSVPVAMPGRGPSALSTP 252
Db 226 VVQDEETDLAP-----SHWAATGDM-----PTYQIRAPTTALPGVVAASPGSL 272

QY 253 TAPHKLGQSLVLTTEERKTLVAEGYPIPTKPLTKSEKALKIRKIKNKISQAQESR 312
Db 273 HSPQOL-----ASEATRKEELRLMKNRAREAR 301

QY 313 RKKEYSMSLEKKEVSCSTENLELRKKVEVLENTNRTLLQLOKLOTLVMGK 364
Db 302 RKKEVYKCLE-----NRVALESSONKTLIEELKALKLDLYCHK 339

RESULT 10

JC4028
activating transcription factor 2 - African clawed frog
N;Alternate names: cyclic AMP-response element-binding protein
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 26-Aug-1999
C;Accession: JC4028
R;Villarreal, X.C.; Richter, J.D.
Gene 153, 225-229, 1995
A;Title: Analysis of ATF2 gene expression during early Xenopus laevis developme nt.
A;Reference number: JC4028; MUID:95180723; PMID:7875593
A;Accession: JC4028
A;Molecule type: mRNA
A;Residues: 1-486 <VIL>
A;Cross-references: GB:U16158; NID:988779; PIDN:AAA69518.1; PID:g710326
C;Comment: This protein is a sequence-specific DNA-binding protein that mediates transcr iption regulation.
Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol ogy
Keywords: leucine zipper; phosphoprotein; transcription regulation
F;329-369/Region: leucine zipper motif
F;329-369/Domain: fos/jun DNA-binding domain homology <FJD>
F;332-364/Region: basic
F;82/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F;102/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 6.1%; Score 162.5; DB 2; Length 486;
Best Local Similarity 21.5%; Pred. NO. 0.03;
Matches 101; Conservative 68; Mismatches 141; Indels 159; Gaps 24;

QY 52 LSDPF-----LSEKSEMEVEPSPTSPAPLIQAE-HSVLSLEERTQSPFTH--AATSD 102
Db 70 LASPPNEFKKASEEDKKPDLPLSLA-TPIIRSKIEPSVVETTHQDSPLPHEPSTSD 128

QY 103 SFNDEVESEKWLSTEPSATIKKEPITEEQPGL-VPSVTLT-----ITAISTPPEK 155
Db 129 QHNQVPL-----AQAQPTSIAVR-----PASLQVNPVLLTSSDSVLIQAVPSP 175

QY 156 RESPLDMNAGDSSCOTL-----IPKIKLPEHVDQFLNFPKESVDQLHLP--- 203
Db 176 SSTVITQAPSTNSSIVTLPGPVLVLLHLPSCQTMFVAI-----PASITSSNVHVA 228

QY 204 -----PTPPSSHSSDSEGLSPNRLHPSLSQAHSVPVAMPGRGPSALST 248

Db 229 PLLRLPLTWVPVPGIPGPSSQPQVQSEAKMRNA-----AWTQOHPV---TNGDTV--- 277

QY 249 SPLLTAPHKLGQSGPLVLTEEKR-----TLVAEGYPIPTK-LPLTKS----- 290

Db 278 -----KGHSFAITQSESRPQSLQQPATSTTETRPSAQPTPQTQSTSGRRRAA 329

QY 291 -----BEKALKIRRKIKNKISQAESRRKKKYMDSLEKKVSCSTENLELRKKVEVLENT 346

Db 330 NEDPDEKRSKIQR---NRAAASRCRQKRWVQVQSEKKADELLISLNGYLQNEVTLRR- 385

QY 347 NRTLQLOKLOTLVMGKVSRTCKLAGTGTGCLMMVVLCPAVAFSGFPQGYGPSATK 406

Db 386 -----EVAQLKQLL--AHKDCPTAMQKKS-----GHSADK 416

QY 407 -----MALPSQHPLEBPYTSVVRNLLIYEHAPLEESSSPASTGE 449

Db 417 DDSSEDISVP-----SSPTEAI-----QHSSVSTSGVSSTSK 450

RESULT 11

TVQJUN
transforming protein jun - Japanese quail
C;Species: Coturnix coturnix japonica (Japanese quail)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 18-Jun-1999
C;Accession: S05963
R;Brun, G.; La Vista, N.; Dangi, J.P.; Castellazzi, M.
Nucleic Acids Res. 17, 6393, 1989
A;Title: Nucleotide sequence of the quail c-jun protooncogene.
A;Reference number: S05963; MUID:89366673; PMID:2505235
A;Accession: S05963
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-313 <BRU>
A;Cross-references: EMBL:X15547; NID:962639; PIDN:CAA33553.1; PID:g62640
C;Genetics: jun
C;Superfamily: jun transforming protein; fos/jun DNA-binding domain homology
C;Keywords: DNA binding; leucine zipper; oncogene; transcription regulation;
F;229-269/Domain: fos/jun DNA-binding domain homology <FJD>
F;262-290/Region: leucine zipper motif

Query Match 6.1%; Score 162; DB 1; Length 313;
Best Local Similarity 22.5%; Pred. No. 0.018;
Matches 72; Conservative 66; Mismatches 128; Indels 54; Gaps 12;

QY 88 EPRTQSPFTHAATSDSFNDEVESEKWLSTEPSATIKKEPITEEQPGLVPSVTLTIT 147
Db 11 EDALNASTAPPESAYGYNNAKVLKQSMNLNSDPASSLK--PHLRKNADIITSPDVGLL 68

QY 148 AISTPFKEESPLDMNAG---GDSSCOTLIPKIKLEPHE--VDQFLNFPKESASVDQLH 201
Db 69 KLASP-ELERLIQSSNGLITTTPTPTQFLCPKNVTDQEGFAEGFVR-----ALAEHL 121

QY 202 LPPTPPS---SHSSSEGLSP-----NPRLHPSLSQAHSVPVRA--MPRPSALS 247

Db 122 NQNTLPSVTSAAQPVSGMVPVSSMAGGSPNTSLH-----SEPPVYANLSNFPNALN 175

QY 248 TSLTLTA-----PHKLOGSP-----LVLTERKRTLVAEGYPIPTKPLTKSEKAL 295

Db 176 SAPNVNANGMVAPOHHNPNQVPHPRLOALKEEFQVPEMPGETPPLSPIDMESQRI 235

QY 296 KKIRRKIKNKISQAESRRKKKEYMDSLEKKVSCSTENLELRKKVEVLENTNRTLLQLO 355

Db 236 KAERKRMNRNIAASKCRKRLERLARLEEKVKTLKAQNSE-----LASTANMLREQVA 288

QY 356 KLOTLVMGKVSRTCKLACTQ 375

Db 289 QLKQKVMNHNVNSGCOLMLTQ 308

RESULT 12

S36101
cAMP response element-binding protein epsilon - mouse

```

Db      229 TG-----EGQKKGRSP----ADKESKRLKLILNRVSAQAARERKAYLI 277
QY      321 SLEKKVSCSTENLRLKKVEVLNTNRTLLQLOQLQTLVMGKVSRCTCKLAGTQTGT 378
        ||::|:||||::||::||::||::||::||::||::||::||::||::||:
Db      271 DLETRVKDLKKNSLKELRSLTLQNEQMRLQILK-----NTTASRGSSNGMT 318

RESULT 14
TFVJVJN
C:Transforming protein jun - avian sarcoma virus 17
C:Species: avian sarcoma virus 17
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Feb-1999
C:Accession: A29870
R;Maki, Y.; Bos, T.J.; Davis, C.; Starbuck, M.; Vogt, P.K.
Proc. Natl. Acad. Sci. U.S.A. 84, 2848-2852, 1987
A>Title: Avian sarcoma virus 17 carries the jun oncogene.
A:Reference number: A29870; MUID:87204225; PMID:3033666
A:Accession: A29870
A:Molecule type: DNA
A:Residues: 1-296 <MA>
C:Comment: This protein is translated as a gag-jun polyprotein.
C:Genetics:
A:Gene: jun
C:Superfamily: jun transforming protein; fos/jun DNA-binding domain homolog
C:Keywords: DNA binding; leucine zipper; oncogene; transcription regulation
F;212-252/Domain: fos/jun DNA-binding domain homology <FJD>
F;245-273/Region: leucine zipper motif

Query Match          63.0%; Score 159.5; DB 1; Length 296;
Best Local Similarity 23.0%; Pred. No. 0.023;
Matches 65; Conservative 45; Mismatches 87; Indels 85; Gaps

QY      117 STEPPSATIKKBPIREQ-----PGLVPVSV-----TL-TITAISPTFEKEPSFLDNAG 150
        ||::|:||||::||::||::||::||::||::||::||::||::||::||:
Db      72 TTPTQTQLCPKNVTDRQGFARGFVRALAEHLNQNTLPSTVTSAAQPVSGMAPVSSMAG 131
QY      166 GDSSCQTLIPKILEPHEVDQFNFSPEAS-----VDQLHLPPTPPSSHSD 213
        ||::|:||||::||::||::||::||::||::||::||::||::||::||:
Db      132 GGSPNTSL---HSEPPVYANLNFNPNALNPYNNRMGVAPQHINPQVPQH-- 184
QY      214 SEGSLSNPLRHPLFSLQAHSVPYRAMRPGPSALSPTSPLLTAHPKLOGSGLPLVITEBKRT 2737
        |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db      185 -----PRLQ--ALKSEPQTVPMPG-----ETPL-----2020
QY      274 LVAGYPIPTKPLPTKSEEKALKIRKIKNKSIAQESRRKKKEYMDLSLEKKVESCSSTEN 333
        |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db      208 -----FPIDMESQBRIKAERKMRNRRIAASKRKRLKLERIARLEEKVKTAKQN 250
QY      334 LELRKVEVLNTNRTLLQLOQLQTLVMGKVSRCTCKLAGTO 375
        |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db      257 SE-----LASTANMLREQVAQLQKVMHNHVGCCQLMLTQ 291

RESULT 15
A37944
CAMP response element-binding protein alpha - mouse
C:Species: Mus musculus (house mouse)
C>Date: 12-Jul-1991 #sequence_revision 12-Jul-1991 #text_change 05-Nov-1991
C:Accession: A37944
R;Foulkes, N.S.; Borrelli, E.; Sassone-Corsi, P.
Cell 64, 739-749, 1991
A>Title: CREM gene: use of alternative DNA-binding domains generates mult
A:Reference number: A37944; MUID:91145994; PMID:1847666
A:Accession: A37944
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-229 <FOU>
A:Cross-references: GB:M60285; NID:g1927716; PIDN:AAA17497.1; PID:g1927719
C:Superfamily: fos/jun DNA-binding domain homology
C:Keywords: alternative splicing; DNA binding; transcription regulation
F;163-206/Domain: fos/jun DNA-binding domain homology <FUD>
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2003, 06:17:30 ; Search time 32 Seconds
(without alignments)
675.286 Million cell updates/sec

Title: US-09-884-566A-2
Perfect score: 2650
Sequence: 1 MEVLESGEQSLQDRKLSL.....LEGNETLKVVLERRVNAVTF 521

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

all number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	445.5	16.8	515	1 CRBA_DROME	P29747 drosophila
2	252	9.5	699	1 ATFB_MOUSE	Q35451 mus musculus
3	233.5	8.4	703	1 AT6B_HUMAN	Q95941 h cyclic-am
4	212	8.0	670	1 AT6A_HUMAN	P18850 homo sapien
5	169	6.4	314	1 APL_SERCA	P54864 serinus can
6	167.5	6.3	741	1 NFL1_MOUSE	Q61985 mus musculus
7	162.5	6.1	341	1 CREM_RAT	Q03061 rattus norv
8	162	6.1	313	1 APL_COTJA	P12981 coturnix co
9	159.5	6.0	296	1 TUUN_AVIS1	P05411 avian sarco
10	158.5	6.0	271	1 ATF1_HUMAN	P18846 homo sapien
11	158	6.0	344	1 CREM_CANFA	P79145 canis famil
12	157.5	5.9	269	1 ATF1_MOUSE	P81269 mus musculus
13	157	5.9	772	1 NFL1_HUMAN	Q14494 h nuclear f
14	156.5	5.9	310	1 APL_CHICK	P18870 gallus gall
15	152.5	5.8	373	1 NFE2_HUMAN	Q18621 homo sapien
16	152	5.7	308	1 JUNB_CYPCA	P18703 cyprinus ca
17	152	5.7	334	1 APL_MOUSE	P05627 mus musculus
18	152	5.7	489	1 YEN5_YEAST	P39970 saccharomyc
19	150.5	5.7	331	1 APL_FIG	P56432 sus scrofa
20	150	5.7	341	1 JUND_MOUSE	P15066 mus musculus
21	150	5.7	341	1 JUND_RAT	P52909 rattus norv
22	150	5.7	605	1 NFL2_HUMAN	Q16236 homo sapien
23	150	5.7	679	1 YIS3_YEAST	P40563 saccharomyc
24	149.5	5.6	325	1 CREB_BOVIN	P27925 bos taurus
25	149.5	5.6	1844	1 POLR_TYMYA	P20128 turnip yell
26	149	5.6	334	1 APL_RAT	P17325 rattus norv
27	149	5.6	494	1 ATF7_HUMAN	P17544 homo sapien
28	148.5	5.6	487	1 ATF2_RAT	Q00969 rattus norv
29	147.5	5.6	341	1 CREM_MOUSE	P27699 mus musculus
30	147.5	5.6	487	1 ATF2_HUMAN	P15336 homo sapien
31	147.5	5.6	487	1 ATF2_MOUSE	P16951 mus musculus
32	146.5	5.5	331	1 APL_HUMAN	P05412 homo sapien
33	146.5	5.5	743	1 TPE3_HUMAN	P19532 homo sapien

Query Match 16.8%; Score 445.5; DB 1; Length 515;

34	146.5	5.5	1462	1 NCO2_MOUSE	Q61026 mus musculus
35	146	5.5	341	1 CREB_RAT	P15337 rattus norv
36	146	5.5	349	1 ATF4_MOUSE	Q06507 mus musculus
37	146	5.5	1232	1 YOO5_CABEL	P34643 caenorhabdi
38	144.5	5.5	597	1 NFL2_MOUSE	Q60795 mus musculus
39	144	5.4	347	1 JUND_HUMAN	P17535 homo sapien
40	144	5.4	746	1 PCAP_HUMAN	Q96rn5 homo sapien
41	142.5	5.4	341	1 CREB_HUMAN	P16220 homo sapien
42	142.5	5.4	341	1 CREB_MOUSE	Q01147 mus musculus
43	142	5.4	289	1 APL_DROME	P18289 drosophila
44	141	5.3	401	1 CPR2_PETCR	Q99090 petroselinu
45	140	5.3	168	1 HVS_ARATH	Q24646 arabidopsis

ALIGNMENTS

RESULT 1
CRBA_DROME
ID CRBA_DROME STANDARD; PRT; 515 AA.
AC P29747;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cyclic-AMP response element binding protein A (Box B binding factor-2)
DE (BBF-2).
GN CRBA OR BBF2.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Oregon-R;
RX MEDLINE=92192458; PubMed=1532159;
RA Abel T., Bhatt R., Maniatis T.;
RT body- and liver-specific regulatory elements."
RL Genes Dev. 6:466-480(1992).
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO FAT BODY-SPECIFIC
CC ENHANCERS OF ALCOHOL DEHYDROGENASE (ADH) AND YOLK PROTEIN GENES.
CC BBF-2 MAY PLAY A ROLE IN FAT BODY GENE EXPRESSION. IT BINDS THE
CC CONSENSUS SEQUENCE 5'-T(A/C)NACGTAN(T/G)C-3'.
CC -!- SUBUNIT: MAY BIND DNA AS HETERODIMERS WITH OTHER BZIP PROTEINS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: IN ALL CELL TYPES EXAMINED.
CC -!- DEVELOPMENTAL STAGE: PRESENT THROUGHOUT DEVELOPMENT.
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X64429; CAA45771.1; -;
CC PIR; S24542; S24542.
CC PIR; A42140; A42140.
CC TRANSFAC; T01603; -;
CC FlyBase; FBgn0004396; Creba.
CC InterPro; IPR004827; TF_bzip.
CC Pfam; PF00170; bzip; 1
CC SMART; SM00338; BRLZ; 1
CC PROSITE; PS00036; BZIP_BASIC; 1.
CC Transcription regulation; Activator; DNA-binding; Nuclear protein.
CC DNA BIND 442 462 BASIC MOTIF.
CC DOMAIN 468 503 LEUCINE-ZIPPER.
CC SEQUENCE 515 AA; 56528 MW; 0808FB9655200223 CRC64;
SQ

Best Local Similarity 33.8%; Pred. No. 8.7e-18;
Matches 142; Conservative 54; Mismatches 117; Indels 107; Gaps 16;

QY 26 ETEALMYH---THFSELDFSONVLQGLSDPFFLSEKSEMEVEPSPTSPAPLQAEHS 82
Db 111 DPKCLTFHVPVTHATP-----ISRLSSNPALN-----TSVADLTRSSGL 149

QY 83 YSL-SSEPTQSPFTHAATSD-----SFNDEVESEKWL-----STPEPS 122
Db 150 QSLQAHQPHGGSSHHVVANLEHFLQPHLYDNCSSVSLRDSGMDPDCSDIEIDE 209

QY 123 ATKKEPTEPPGGLVPS-----VLTITATSTP---FEKESPLDMNAGDSSQ 171
Db 210 SAIKDEPMSPDSCPASPTSQSSSQHQLSLNLAHQSEMLFEPKHCGLLTFASSNSN-N 268

QY 172 TLIPK-----IKLEPHEVDQFLNFSPEASVDQLHLPPTPPPS-HSSDSE 215
Db 269 SLIKSQORHEQILQDNLMLAKMEIKSEKQSTNSSSGKSHAHGYGPTTPPSLPSDDSE 328

216 GSLSPN-----PRL-----HPSLSQAHSPVR-----AMRGPSTALST-- 248
329 GNLSPHFLRHCRTPTQFPFLMFIQPAVSHPYGSATAASTRSSGASASGTSSTVT 388

QY 249 -----SPLTAPHKLGSGPLVLTETEEKRTLVAGYPIPTKPLTKSBEKALKIRK 301
Db 389 TTRQPTHTPLISSQPK-GSTGTLTLLTEEEKRTLVAGYPIQKPLTKAEKSLKIRK 447

QY 302 IKNKISAESRRKKYMPDSLEKKVSCSTENLELRKVEVLENTNRTLLQQLQTLV 361
Db 448 IKNKISAESRRKKYMDQLERRVBLVTENHYKRLGLEBTNANLLSQLKLQALV 507

RESULT 2
ATPB MOUSE STANDARD; PRT; 699 AA.
AC O35451,
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Cyclic-AMP-dependent transcription factor ATF-6 beta (Activating transcription factor 6 beta) (ATF6-beta) (cAMP responsive element binding protein-like 1) (cAMP response element binding protein-related protein) (creb-rp).
DE CREBL1.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]_TaxID=10090;
SEQUENCE FROM N.A.
RA Rowen L., Mahairas G., Qin S., Ahearn M.E., Dankers C., Lasky S., Loretz C., Schmidt S., Tipton S., Traicoff R., Zackrone K., Hood L.;
RA "Sequence of the mouse major histocompatibility locus class III region."
RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RP Straussberg R.;
RC TISSUE=Kidney;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 654-699 FROM N.A.
RC STRAIN=C57BL/6 X CBA;
RX MEDLINE=99014231; PubMed=9795100;
RA Ikuta T., Sogawa N., Ariga H., Ikemura T., Matsumoto K.;
RT "Structural analysis of mouse tenascin-X: evolutionary aspects of reduplication of FNIII repeats in the tenascin gene family."
RL Gene 217:11-13 (1998).
CC -1- FUNCTION: TRANSCRIPTIONAL FACTOR THAT ACTS IN THE UNFOLDED PROTEIN RESPONSE (UPR) PATHWAY BY ACTIVATING UPR TARGET GENES INDUCED DURING ER STRESS. BINDS ON THE 5'-CCAC[GA]-3' HALF OF THE ER STRESS RESPONSE ELEMENT (ERSE) (5'-CCAAT-N9-CCAC[GA]-3') WHEN NF-Y IS BOUND TO ERSE (BY SIMILARITY).

CC -1- SUBUNIT: HOMODIMER AND HETERODIMER WITH ATF6-ALPHA. THE DIMER INTERACTS WITH THE NUCLEAR TRANSCRIPTION FACTOR Y (NF-Y) TRIMER THROUGH DIRECT BINDING TO NF-Y SUBUNIT C (NF-YC) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein in the endoplasmic reticulum. Under ER stress the cleaved N-terminal cytoplasmic domain translocates into the nucleus (By similarity).
CC -1- PTM: DURING UNFOLDED PROTEIN RESPONSE AN APPROXIMATIVE 60 KDA FRAGMENT CONTAINING THE CYTOPLASMIC TRANSCRIPTION FACTOR DOMAIN IS RELEASED BY PROTEOLYSIS. THE CLEAVAGE IS PROBABLY PERFORMED SEQUENTIALLY BY SITE-1 AND SITE-2 PROTEASE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. ATF SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL: AF030001; AAB82014.1; -;
DR EMBL: BC013534; AAH33534.1; -;
DR EMBL: AB010266; BAA24435.1; -;
DR MGI: 105121; Creb1.
DR InterPro: IPR004827; TF_bZIP.
DR Pfam: PF00170; bZIP; 1;
DR SMART: SM00338; BRLZ; 1;
DR PROSITE: PS00036; BZIP_BASIC; 1;
DR Transcription regulation; DNA-binding; Activator;
KW Unfolded protein response; Nuclear protein; Endoplasmic reticulum;
KW Transmembrane; Signal-anchor; Glycoprotein.
FT DOMAIN 1 393 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 394 414 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).
FT DOMAIN 415 699 LUMENAL (POTENTIAL).
FT DNA_BIND 324 344 BASIC DOMAIN.
FT DOMAIN 364 385 LEUCINE-ZIPPER.
FT DOMAIN 93 102 POLY-SER.
FT SITE 407 407 IMPORTANT FOR CLEAVAGE BY PS2 (BY SIMILARITY).
FT SITE 410 410 IMPORTANT FOR CLEAVAGE BY PS2 (BY SIMILARITY).
FT SITE 437 438 CLEAVAGE (BY PSI) (BY SIMILARITY).
FT CARBOHYD 473 473 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 502 502 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 699 AA; 76007 MW; 6EBE22386D63868B CRC64;

Query Match 9.5%; Score 252; DB 1; Length 699;
Best Local Similarity 23.5%; Pred. No. 5.8e-07;
Matches 127; Conservative 69; Mismatches 204; Indels 140; Gaps 21;

QY 14 WDRKL-SELSPGCTEALMVHTFSELLDFEQNLVQLLS-----DPLSEKSESMEV 66
Db 27 WDSLTLGDLDEVAERQQLFRCEQDVDFDSSLDVGMVSPPEPPWDPLPIFDFLQVKS 86

QY 67 EPSPTSPAPLIQAEHSYLSSEPTQ-----SPF----- 95
Db 87 EPSPPCSSSSLSSESH-LSTETPPSQVGVGVHLVKMESLAPPLCLLGDGDDPASPPETVQ 145

QY 96 -THAATSDSPND-----EEVESEKWLSTPEPSATIKKEPTE-----EQPPGL 138
Db 146 ITVGSASDDLSDIQTKLEPASPSVSHSEASLSADSPPQFFIGEEVLEVKTSPPSPGC 205

QY 139 -----VPSVTLTITATSTPEKESPLDMNAGDSSQCTLLPKIKLEPHEVDQFLNFSPE 194
Db 206 LMDVDPASSLQAVQIS-----MGSPDSS-SGKAPATRKPP-----LQPKP 245

QY 195 ASVDQLHLPTPTPSSHSSDEGSLSP---NRLHPFSLSQAHSPVRAMPGPSALSTSL 251
Db 246 VLTTV---VPPPRAGTSAVLLQLPLVQPAVSPVVLIQ--GAIRVQPEGPAPAPRP- 299

QY 252 LTAPHKLGSGPLVTESEKRTVAEGYPIPTKPLTKSEKALKIRKIKKKAQES 311
 Db 300 -----ERKSIVPA--PMPGNSCPPEVDKLLKRRQRMKIKRESACQS 339
 QY 312 RKKKEYMDSLEKKVESSTENLELRKKEVELENTNRTLLQQLQKLQTLVMGKVSRTCKL 371
 Db 340 RKKKEYLQGLLEARQAVLADNQLRRRENAALRRLEALLAENSGLK---LG----- 388
 QY 372 AGTQTGTCMLVVLCAVAFGSF-PQGGPYPSATKMPALPSQHPSEPYTASVRSRMLL 430
 Db 389 SGNRKVCVMVFLFLIAFNGFVSVISEPPAPMSPMSRREPRP-----QRHLL 437
 QY 431 IVEEHAPLESSSPASTGELGWDGSGSLLRAS-----SGLALPEVDLPHELIISNE 482
 Db 438 GFSEPGPAHMEPLREAAQSGEQSPSAGRPFRNLTAFFPGAKELLRLDLDQLFLSSD 497

RESULT 3

HUMAN
 AT6B HUMAN STANDARD; PRT; 703 AA.
 Q99941; Q14343; Q14345; Q13269; Q9NPL0; Q99635; Q99637; Q9H3V9;
 Q9H3W1;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cyclic-AMP-dependent transcription factor ATF-6 beta (Activating
 DE transcription factor 6 beta) (ATF6-beta) (cAMP responsive element
 DE binding protein-like 1) (cAMP response element binding protein-related
 DE protein) (Creb-fp) (G13 protein).
 GN CREB1 OR G13.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Tonsil;
 RX MEDLINE=96163866; PubMed=8586413;
 RA Min J., Shukla H., Kozono H., Bronson S.K., Weisman S.M.,
 RA Chaplin D.D.;
 RT "A novel Creb family gene telomeric of HLA-DRA in the HLA complex.";
 RL Genomics 30:149-156(1995).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=97024424; PubMed=8870652;
 RA Khanna A., Campbell R.D.;
 RA "The gene G13 in the class III region of the human MHC encodes a
 potential DNA-binding protein.";
 RL Biochem. J. 319:81-89(1996).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Rowen L., Danks C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
 RA Banta A., Schwartzell S., Smith T.M., Spies T., Hood L.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=97081760; PubMed=8923003;
 RA Speck M., Barry F., Miller W.L.;
 RT "Alternate promoters and alternate splicing of human tenascin-X, a
 RT gene with 5' and 3' ends buried in other genes.";
 RL Hum. Mol. Genet. 15:1749-1758(1996).
 RN [5]
 RP SEQUENCE OF 158-700 FROM N.A.
 RA Barlow K.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP CHARACTERIZATION.
 RX MEDLINE=21157024; PubMed=11256944;
 RA Haze K., Okada T., Yoshida H., Yanagi H., Yura T., Negishi M.,
 RA Mori K.;
 RT "Identification of the G13 (cAMP-response-element-binding
 RT protein-related protein) gene product related to activating

transcription factor 6 as a transcriptional activator of the
 mammalian unfolded protein response.";
 RL Biochem. J. 355:19-28(2001).
 CC -!- FUNCTION: TRANSCRIPTIONAL FACTOR THAT ACTS IN THE UNFOLDED PROTEIN
 CC RESPONSE (UPR) PATHWAY BY ACTIVATING UPR TARGET GENES INDUCED
 CC DURING ER STRESS. BINDS DNA ON THE 5'-CCAC[GA]-3' HALF OF THE ER
 CC STRESS RESPONSE ELEMENT (ERSE) (5'-CCAAT-N9-CCAC[GA]-3') WHEN NF-Y
 CC IS BOUND TO ERSE.
 CC -!- SUBUNIT: HOMODIMER AND HETERODIMER WITH ATF6-ALPHA. THE DIMER
 CC INTERACTS WITH THE NUCLEAR TRANSCRIPTION FACTOR Y (NF-Y) TRIMER
 CC THROUGH DIRECT BINDING TO NF-Y SUBUNIT C (NF-YC).
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein in the endoplasmic
 CC reticulum. Under ER stress the cleaved N-terminal cytoplasmic
 CC domain translocates into the nucleus.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 AND 2 (SHOWN HERE); MAY BE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: UBIQUITOUS.
 CC -!- DOMAIN: THE BASIC DOMAIN FUNCTIONS AS A NUCLEAR LOCALIZATION
 CC SIGNAL.
 CC -!- DOMAIN: THE BASIC LEUCINE-ZIPPER DOMAIN IS SUFFICIENT FOR
 CC ASSOCIATION WITH THE NF-Y TRIMER AND BINDING TO ERSE.
 CC -!- PTM: N-GLYCOSYLATED.
 CC -!- DURING UNFOLDED PROTEIN RESPONSE AN APPROXIMATIVE 60 KDA
 CC FRAGMENT CONTAINING THE CYTOPLASMIC TRANSCRIPTION FACTOR DOMAIN IS
 CC RELEASED BY PROTEOLYSIS. THE CLEAVAGE IS PROBABLY PERFORMED
 CC SEQUENTIALLY BY SITE-1 AND SITE-2 PROTEASES.
 CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. ATF SUBFAMILY.
 CC -!- CAUTION: SOME REF.4 SEQUENCES DIFFER FROM THAT SHOWN DUE TO A
 CC FRAMESHIFT IN POSITION 600.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U31903; AAA97438.1; -;
 CC EMBL; X98053; CAA66663.1; -;
 CC EMBL; X98054; CAA66664.1; -;
 CC EMBL; U89337; AAB47487.1; -;
 CC EMBL; U52694; AAG14900.1; -;
 CC EMBL; U52696; AAC50888.1; ALT FRAME.
 CC EMBL; U52693; AAG14898.1; ALT FRAME.
 CC EMBL; U52701; AAC50883.1; ALT SEQ.
 CC EMBL; AL049547; CAB89295.1; -;
 CC Genew; HGNC:2349; CREB1.
 CC InterPro; IPR004827; TF_bZIP.
 CC MIM; 600984; -;
 CC Pfam; PF00170; bZIP; 3.
 CC SMART; SM00338; BRLZ; 1.
 CC PROSITE; PS00036; BZIP_BASIC; 1.
 CC Transcription regulation; DNA-binding; Activator;
 CC Unfolded protein response; Nuclear protein; Endoplasmic reticulum;
 CC Transmembrane; Signal-anchor; Glycoprotein; Alternative splicing.
 CC DOMAIN 1 396
 CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 CC (POTENTIAL).
 CC TRANSMEM 397 417
 CC LUMENAL (POTENTIAL).
 CC DOMAIN 418 703
 CC TRANSCRIPTION ACTIVATION.
 CC DOMAIN 1 86
 CC DNA_BIND 327 347
 CC BASIC DOMAIN.
 CC DOMAIN 367 388
 CC LEUCINE-ZIPPER.
 CC DOMAIN 96 105
 CC POLY-SER.
 CC SITE 410 410
 CC IMPORTANT FOR CLEAVAGE BY PS2 (BY
 CC SIMILARITY).
 CC SITE 413 413
 CC IMPORTANT FOR CLEAVAGE BY PS2 (BY
 CC SIMILARITY).
 CC SITE 440 441
 CC CLEAVAGE BY PS1 (BY SIMILARITY).
 CC SITE 476 476
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 505 505
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 610 610
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SITE 627 627
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 627

```
FT CARBOHYD 676 676 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSPLIC 28 31 GLQN -> D (IN ISOFORM 1).
FT CONFLICT 3 3 E -> D (IN REF. 2).
FT CONFLICT 329 330 QQ -> HE (IN REF. 2).
FT CONFLICT 520 520 V -> D (IN REF. 4; AAC50888).
FT CONFLICT 600 600 D -> G (IN REF. 1).
SQ SEQUENCE 703 AA; 76709 MW; B04C7B2337D83F82 CRC64;

Query Match 8.4%; Score 223.5; DB 1; Length 703;
Best Local Similarity 23.5%; Pred. NO. 2.2e-05;
Matches 135; Conservative 78; Mismatches 216; Indels 145; Gaps 24;

Qy 18 LSEUSEP-----GETEALMTHFSELLDFESQNLQQLSDPFL-----SEKS 61
Db 7 LSEADPTFRFTDNLSPEDWGLQNSTLYSG-LDEVAEEQT-QLFRCPEDQVPPFGSSLD 64
Qy 62 ESMEVEPPTSPAPLIQAHSYLSLEERTQSPFTTAATSDSFNDEEVESEKWLSTTFPP 121
Db 65 VQMDVSPS-EPPWELLPDPDLOVKSEP--SSPCSSSLU-----SESRLLSTEPS 112
Qy 122 SATIKKEPTTEQPPLVPSVTLTITATSTPEKESPLDMNAGDSSQTLIPKIKLEP 181
Db 113 SEALGVGEVLHVKTSLAPPLCLGDDPTSSPTQINVIPTSDSDSDVQT-----KIEP 167
Qy 182 HEVDQFLNFPKESVDQ-----LHLPPTP-----P 207
Db 168 VSPCSSVNSEASLLSADSSQAFIGEVLEVKTESLSPSGCLLDVDPAPSLGAVQISMGP 227
Qy 208 SSHSDSEGLSPNRLH--PFLSQAHSPVRAMPRG-----PSALSTSPLLTAP 255
Db 228 SLDGSGKALPRKPPLOPKPVLLITVPMPSRAVPSTTVLLQSLVQPPVPFVLIQA 287
Qy 256 HKLQSGPL-VLTBEKRTLVAGVPPIPTKLPTKSEKALKIRKIRKIKKISAOBSRRK 314
Db 288 IRVQEPGAPSLPRPKKIVPA--PMPGNSCPPEVDKLLKQKQRMKNRESACQSRK 345
Qy 315 KKEYMDSLEKKVESCSTENLEIRKKEVELENTNRTLLQQLKLTVMGKVKSTCKLAGT 374
Db 346 KKEYLQGLARQAVLADNQQLRRENAALRRLEALLAENSELK--LG-----SGN 394
Qy 375 QTGTCLMVVLCFAVAFSGFFQGYGYPSPATKALPSQHPLS-----PPTASVVRSN 428
Db 395 RKVCVMVFLFLFTAFNFG-----PVSISEPPSPAPISPMKMGEPQP-----RRH 438
Qy 429 LTIYEHAPLESSESSPASTGELGMDRGSLLRASSGLEALPE---VDLPFLFLISNETSL 485
Db 439 LIGFSQEPVQ-----GVEPLQSGQKPEQPSPTDQPSF--SNLTAF 480
Qy 486 ----EKSVLLELQHLVSS---KLEGNETLKVVE 512
Db 481 PGGAKELLRLDLQFLSLDCRHFNRFTESURLAD 514

RESULT 4
AT6A HUMAN STANDARD; PRT; 670 AA.
AC P18850; O15139; Q9UEC9;
DT 01-NOV-1990 (Rel. 16, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cyclic-AMP-dependent transcription factor ATF-6 alpha (Activating
DE transcription factor 6 alpha) (ATF6-alpha).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervical carcinoma;
RX MEDLINE=97415575; Pubmed=9271374;
RA Zhu C., Johansen F.E., Prywes R.;
RA "Interaction of ATF6 and serum response factor.";
```

```
RL Mol. Cell. Biol. 17:4957-4965(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervical carcinoma;
RX MEDLINE=9057946; Pubmed=9837962;
RA Yoshida H., Haze K., Yanagi H., Yura T., Mori K.;
RA "Identification of the cis-acting endoplasmic reticulum stress
RT response element responsible for transcriptional induction of
RT mammalian glucose-regulated proteins; involvement of basic-leucine
RT zipper transcription factors.";
RL J. Biol. Chem. 273:33741-33749(1998).
RN [3]
RP SEQUENCE OF 302-369 FROM N.A.
RX MEDLINE=90195187; Pubmed=2516827;
RA Hai T., Liu F., Coukos W.J., Green M.R.;
RT "Transcription factor ATF CDNA clones: an extensive family of leucine
RT zipper proteins able to selectively form DNA-binding heterodimers.";
RL Genes Dev. 3:2083-2090(1989).
RN [4]
RP ERRATUM.
RA Hai T., Liu F., Coukos W.J., Green M.R.;
RL Genes Dev. 4:682-682(1990).
RN [5]
RP SEQUENCE OF 1-202 FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP CHARACTERIZATION.
RX MEDLINE=20032174; Pubmed=10564271;
RA Haze K., Yoshida H., Yanagi H., Yura T., Mori K.;
RT "Mammalian transcription factor ATF6 is synthesized as a transmembrane
RT protein and activated by proteolysis in response to endoplasmic
RT reticulum stress.";
RL Mol. Biol. Cell. 10:3787-3799(1999).
RN [7]
RP PROCESSING BY PS1 AND PS2, AND MUTAGENESIS OF ARG-415; ARG-416;
RP LEU-419; ASN-391 AND PRO-394.
RX MEDLINE=2111045; Pubmed=11163209;
RA Ye J., Rawson R.B., Komuro R., Chen X., Dave U.P., Prywes R.,
RA Brown M.S., Goldstein J.L.;
RT "ER stress induces cleavage of membrane-bound ATF6 by the same
RT proteases that process SREBPs.";
RL Mol. Cell. 6:1355-1364(2000).
RN [8]
RP FUNCTION.
RX MEDLINE=21106370; Pubmed=11158310;
RA Yoshida H., Okada T., Haze K., Yanagi H., Yura T., Negishi M.,
RA Mori K.;
RT "Endoplasmic reticulum stress-induced formation of transcription
RT factor complex ERSPF including NF-Y (CBF) and activating transcription
RT factors 6alpha and 6beta that activates the mammalian unfolded
RT protein response.";
RL Mol. Cell. Biol. 21:1239-1248(2001).
RN [9]
RP REVIEW.
RX MEDLINE=21376119; Pubmed=11483355;
RA Hai T., Hartman M.G.;
RT "The molecular biology and nomenclature of the activating
RT transcription factor/CAMP responsive element binding family of
RT transcription factors: activating transcription factor proteins and
RT homeostasis.";
RL Gene 273:1-11(2001).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT ACTS DURING ENDOPLASMIC
CC RETICULUM STRESS BY ACTIVATING UNFOLDED PROTEIN RESPONSE TARGET
CC GENES. BINDS DNA ON THE 5'-CCAC[GA]-3' HALF OF THE ER STRESS
CC RESPONSE ELEMENT (ERSE) (5'-CCAAT-N9-CCAC[GA]-3') AND OF ERSE II
CC (5'-ATTGG-N-CCACG-3'). BINDING TO ERSE REQUIRES BINDING OF NF-Y TO
CC ERSE. COULD ALSO BE INVOLVED IN ACTIVATION OF TRANSCRIPTION BY THE
CC SERUM RESPONSE FACTOR.
CC -!- SUBUNIT: HOMODIMER AND HETERODIMER WITH ATF6-BETA. THE DIMER
CC INTERACTS WITH THE NUCLEAR TRANSCRIPTION FACTOR Y (NF-Y) TRIMER
CC THROUGH DIRECT BINDING TO NF-Y SUBUNIT C (NF-YC). INTERACTS ALSO
```

WITH THE TRANSCRIPTION FACTORS GTF2I, YY1 AND SRF.
 -!- SUBCELLULAR LOCATION: Type II membrane protein in the endoplasmic reticulum. Under ER stress the cleaved N-terminal cytoplasmic domain translocates into the nucleus.
 -!- TISSUE SPECIFICITY: UBIQUITOUS.
 -!- DOMAIN: THE BASIC DOMAIN FUNCTIONS AS A NUCLEAR LOCALIZATION SIGNAL.
 -!- DOMAIN: THE BASIC LEUCINE-ZIPPER DOMAIN IS SUFFICIENT FOR ASSOCIATION WITH THE NF-Y TRIMER AND BINDING TO ERSE.
 -!- PTM: DURING UNFOLDED PROTEIN RESPONSE AN APPROXIMATIVE 50 KDA FRAGMENT CONTAINING THE CYTOPLASMIC TRANSCRIPTION FACTOR DOMAIN IS RELEASED BY PROTEOLYSIS. THE CLEAVAGE SEEMS TO BE PERFORMED SEQUENTIALLY BY SITE-1 AND SITE-2 PROTEASES.
 -!- PTM: N-GLYCOSYLATED.
 -!- PTM: PHOSPHORYLATED IN VITRO BY MAPK14/P38MAPK.
 -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. ATF SUBFAMILY.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

EMBL; AF005887; AAB64434.1; --
 EMBL; AB015856; BAA34722.1; --
 EMBL; BC014969; AAH14969.1; --
 PIR; F34223; F34223.
 HSSP; P03069; 1YSA.
 TRANSPAC; T04742; --
 Genew; HGNC:791; ATF6.
 MIM; 605537; --
 InterPro; IPR004827; TF_bzip.
 Pfam; PF00170; bzip; 1.
 SMART; SM00338; BRL2; 1.
 PROSITE; PS00036; BZIP_BASIC; 1.
 Transcription regulation; DNA-binding; Activator;
 Unfolded protein response; Nuclear protein; Endoplasmic reticulum;
 Transmembrane; Signal-anchor; Glycoprotein; Phosphorylation.
 DOMAIN 1 377
 TRANSMEM 378 398
 (POTENTIAL)
 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 (POTENTIAL)
 LUMENAL (POTENTIAL).
 TRANSCRIPTION ACTIVATION.
 BASIC MOTIF.
 LEUCINE-ZIPPER.
 POLY-SER.
 POLY-LYS.
 POLY-PRO.
 CLEAVAGE (BY P51) (BY SIMILARITY).
 N-LINKED (GLCNAC. .) (POTENTIAL).
 N-LINKED (GLCNAC. .) (POTENTIAL).
 N-LINKED (GLCNAC. .) (POTENTIAL).
 N-LINKED (GLCNAC. .) (POTENTIAL).
 N->F: LOSS OF PROTEOLYTIC CLEAVAGE; WHEN ASSOCIATED WITH L-394.
 P->L: LOSS OF PROTEOLYTIC CLEAVAGE; WHEN ASSOCIATED WITH F-391.
 RR->AA: REDUCES PROTEOLYTIC CLEAVAGE.
 L->V: REDUCES PROTEOLYTIC CLEAVAGE.
 L->M (IN REF. 2).
 L->V (IN REF. 4).
 P->A (IN REF. 4).
 P->P (IN REF. 4).
 N->I (IN REF. 4).
 VPAA->IPPO (IN REF. 4).
 L->I (IN REF. 3).
 T->R (IN REF. 3).
 NORL->LRNS (IN REF. 3).
 G->S (IN REF. 2).
 VV->AL (IN REF. 2).
 670 AA; 74566 MW; CEA70B72FE8954C1 CRC64;

Query Match 8.0%; Score 212; DB 1; Length 670;
 Best Local Similarity 22.2%; Pred. No. 8.8e-05;
 Matches 121; Conservative 79; Mismatches 177; Indels 168; Gaps 23;

QY 14 WDRKL-SELSEPFGETEALM-----YHTFSELDFEQNVGLQLSDPFLSE-----KS 61
 DB 27 WDSALFAELGVFTDDELQLEAANETENNFDNL--DFDLDL-----PWESDIWDINN 78
 QY 62 ESMEVEPPTSPAPLQAEHSYLSSEERTQSPFTHAATSDSFNDEEVESEKWIYST--- 118
 DB 79 QICTVKDIKAEFQPLSPASSYSVS-SPR-----SVDSYSTOHVPELDLSSSSQ 128
 QY 119 -----EFPSATIKKEPITEEQPGLVPSVTLTITAIISTPFKEEESPLDMNAGDSSC 170
 DB 129 MSPSLYGENSNSLSSPELKEDKP-----VTSRNKTEN----- 163
 QY 171 QTLIPKIKLEPHEVDQFLNFPKESVDQLHLPTPPPSHSDSEGS---LSPNRLHPF 227
 DB 164 -GLTPKKI---QVNSKPSIQPK-----PLLLPAAPKQTNSVPAKTIITQVTLMLPL 214
 QY 228 SLSQAHSPVRMPGSPALSPTSPLTAPHKLGSGPLVLTTEEKKTLVAEGY----- 279
 DB 215 AKQQ---PIISLQAPPTKGTQ-TVLLSQPTVVLQAPGVLPSAQPVLAAGGVTQLPNHV 270
 QY 280 -----PIPTKPLPTK-----SEKALKKIRKIKNKISAQESREKKKEY 318
 DB 271 NVVPAPSANSVPNGKLSVTGLVOSTMNVGSDIAVLRQQRQMKIKNRSAOSRKKKKEY 330
 QY 319 MDSLEKVKVESCSTENLELRKKEVLENTNRTLLQLOKLQTLVMGKVRTCKLACTGTGT 378
 DB 331 MGLRLKALKAALSENEQLKKE-----NGTLKRLQDEVVS-----ENQRLKVPSPKRV 378
 QY 379 CLMVVLCFAVAFSGFFQGYGP---YPSATKVALPSQHPLSE----- 417
 DB 379 VCVMLVLAFIIL-----NYGPMLEQDSRRMNPVSGPANORRHLLGFSAKEAQTSDG 432
 QY 418 -----PYTASVVRNLLIYEEHAPLEESSSPAS-----TCGLCGWDRGSSL 459
 DB 433 IIQKNSYRHYDSVNDKALMVLTEPELLYIPPPCQPLINTESLRLNLHRLGWVHRHEV 492
 QY 460 LRASS 464
 DB 493 ERTKS 497

RESULT 5
 API_SERCA
 ID API_SERCA STANDARD; PRT; 314 AA.
 AC P54864;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Transcription factor AP-1 (Proto-oncogene c-jun).
 GN JUN.
 OS Serinus canaria (Canary).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Passeriformes; Passeroides;
 OC Fringillidae; Carduelinae; Serinus.
 OX NCBI_TaxID=9135;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wasserschlag; TISSUE=Brain;
 RX MEDLINE=95206115; PubMed=7898314;
 RA Nastiuk K.L., Mello C.V., George J.M., Clayton D.F.;
 RT "Immediate-early gene responses in the avian song control system:
 RL cloning and expression analysis of the canary c-jun CDNA";
 RL Brain Res. Mol. Brain Res. 27:299-309(1994).
 CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS AND RECOGNIZE THE
 CC ENHANCER DNA SEQUENCE: TGA(C/G)TCA.
 CC -!- SUBUNIT: INTERACTS WITH C-FOS TO FORM A DIMER.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. JUN SUBFAMILY.


```
Db 251 ----PTYQIRAPTALPOGVMAASPGSLHSPQOL----- 281
Qy 284 KLPLTKSEBALKTKRKIKNKISAQESRRKKKYWDSLEKKVESCSTENLELRKKVEVL 343
Db 282 -----AEBATRKRLRLMKNEARECRKKKEYVKCLE-----NRVAVL 321
Qy 344 EYNTNRLLOQLKLOTLVMGK 364
Db 322 ENQNTLIEELKALKDLYCHK 342

RESULT 12
ATF1_MOUSE
ID ATF1_MOUSE STANDARD; PRT; 269 AA.
AC P81369;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cyclic-AMP-dependent transcription factor ATF-1 (Activating
transcription factor 1) (TCF-ATF1).
ATF1.
Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=92176656; PubMed=1531847;
RA Lei M.-R.; Chung C.-S.; Liou M.-L.; Wu M.; Li W.-F.; Hsueh Y.-P.;
RA Lai M.-Z.;
RT "Isolation and characterization of nuclear proteins that bind to T
cell receptor V beta decamer motif.";
RL J. Immunol. 148:1906-1912(1992).
CC -!- FUNCTION: THIS PROTEIN BINDS THE CAMP RESPONSE ELEMENT (CRE)
(CONSENSUS: 5'GTGACGT(A/C)(A/G)-3'), A SEQUENCE PRESENT IN MANY
VIRAL AND CELLULAR PROMOTERS. BINDS TO THE TAX-RESPONSIVE ELEMENT
(TRE) OF HTLV-I.
CC -!- SUBUNIT: BINDS DNA AS A DIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. ATF SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL; M63725; AAA40395.1; -.
MGD; MGI:1298366; Atf1.
DR InterPro; IPR001630; Leuzip_CREB.
DR InterPro; IPR004827; TF_bZIP.
DR InterPro; IPR003102; pKID.
DR Pfam; PF00170; bZIP; 1.
DR Pfam; PF02173; pKID; 1.
DR PRINTS; PR00041; LEUZIPPRCREB.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW Transcription regulation; DNA-binding; Activator; Nuclear protein.
FT DNA_BIND 214 233 BASIC MOTIF.
FT DOMAIN 239 260 LEUCINE-ZIPPER.
SQ SEQUENCE 269 AA; 29237 MW; 9885265159D64A0C CRC64;

Query Match 5.9%; Score 157.5; DB 1; Length 269;
Best Local Similarity 24.7%; Pred. No. 0.028;
Matches 81; Conservative 47; Mismatches 115; Indels 85; Gaps 16;

Qy 58 QSKSSENEV-EPSPTSPAPLI-QAEHSY-SLSEPRTPQSPFTHAATSDSFNDEEVESEK 113
Db 4 SHKSNTTETASQPGSTVAGPHVSIVHQVSLSESESSQD-----SSDSIG-----SSQK 53
Qy 114 WY-LSHPPS-----ATIKKEPITEEPPGLVPSVTLITAITSPFPKESPLDMNAGGD 167
```

```
Db 54 AHGILARRPSYRKILKDLSSDTRGRKGEENPSSIS-AITSMSP-----ADIYQTSSQ 107
Qy 168 SSCQTLIPKIKLEPHEVQDFLNFSPEASVDOLH-LPTTPSS-----HSDSEG 216
Db 108 -----YIAIAPNGALQLA--SPSTDGVQALQTLTWTNSSSTQOQTILQVAQTSQGO 157
Qy 217 SLSPNRLHPFSLSAHSPVRAMPGRPSALSTPLTAPHKLQSGPLVLTTEEKRTLVA 276
Db 158 ILVPSNQV-----VVQTASGDMQTYQIRTPSATSLPQTVMWTS----- 196
Qy 277 EGYPIPTKLPLTKSEBALKTKRKIKNKISAQESRRKKKYWDSLEKKVESCSTENLEL 336
Db 197 ---PVTLASQTTKDDPQLRREIRLMKNREARECRKKKEYVKCLE----- 240
Qy 337 RKKEVLENTNRLLOQLKLOTLVMGK 364
Db 241 -NRVAVLENQNTLIEELKALKDLYCHK 267

RESULT 13
NFL1_HUMAN
ID NFL1_HUMAN STANDARD; PRT; 772 AA.
AC Q14494; Q12877;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nuclear factor erythroid 2 related factor 1 (NF-E2 related factor 1)
(NFE2-related factor 1) (Nuclear factor, erythroid derived 2, like 1)
(DE Transcription factor 1) (Transcription factor HBZ17) (Transcription
factor LCR-F1) (Locus control region-factor 1).
GN NFE2L1 OR NRF1 OR TCF11 OR HBZ17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=95095252; PubMed=8001966;
RA Luna L.; Johnsen O.; Skartlien A.H.; Pedetour F.; Turc-Carel C.;
RA Prydz H.; Kolatse A.-B.;
RT "Molecular cloning of a putative novel human bZIP transcription
factor on chromosome 17q22.";
RL Genomics 22:553-562(1994).
RN [2]
SEQUENCE OF 326-772 FROM N.A.
RP MEDLINE=94310069; PubMed=8036168;
RA Caterina J.J.; Donze D.; Sun C.W.; Ciavatta D.J.; Townes T.M.;
RA "Cloning and functional characterization of LCR-F1: a bZIP
transcription factor that activates erythroid-specific, human globin
gene expression.";
RL Nucleic Acids Res. 22:2383-2391(1994).
CC -!- FUNCTION: ACTIVATES ERYTHROID-SPECIFIC, GLOBIN GENE EXPRESSION.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL; X77366; CAA54555.1; -.
DR EMBL; U08853; AAA20466.1; -.
DR HSP; P34707; LSKN.
DR Genew; HGNC:7781; NFE2L1.
DR MIM; 163260; -.
DR InterPro; IPR004827; TF_bZIP.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
```

```
FT DOMAIN 125 288 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 496 517 POLY-SER.
FT DNA_BIND 659 674 BASIC MOTIF.
FT DOMAIN 682 704 LEUCINE-ZIPPER.
SQ SEQUENCE 772 AA; 84703 MW; C868807C6046BEFS CRC64;

Query Match
Best Local Similarity 5.9%; Score 157; DB 1; Length 772;
Matches 80; Conservative 62; Mismatches 135; Indels 66; Gaps 13;

Qy 28 EALMYTHFSEL-LDFSONVQLLDLSEKESMEVEPSPTSPAPLQAEHSYSL 86
Db 444 EAMLDEISLMDLAIEGFPVQASGLEEFDSGLSLD---SSHSPSLSSSESSSS 500
Qy 87 EBPRTQSPFTHAATSDSFNDEEVESEKWLSTEPSATIKKEPIEEOQPGVPSVTLTI 146
Db 501 SSSSSSSSSASSSSSSSEGAAGV---YSS-----DSETLDLEAGAV----- 542
Qy 147 TAISTPFKEEPLDMDAGDSSCOTLPIKLEPHEVDQFNFSPKEASVDOLHLPPTP 206
Db 543 -----GYQPEYKFCRMSYQDPAQLSCLPYLEHVGH--NHTYNAP--SALDSADLPPP- 592
Qy 207 PSSHSDSEGSLSNPRLHPFSLQAHSPVRAMPGRPSALSTSPLLTAPHKLQSGGLVL 266
Db 593 ----SALKKGSKEQADFLDKQMSRDEHRAAM---KIPFTNDKIINLP----- 634
Qy 267 TEEKRTLVAGYPIPTKPLTKSEKALKIRKIKNKISAEQRRKKKEYMDSLEKKV 326
Db 635 -----VEEFNELLSKYQLSEAGLSLIRRRGKMKVAAQNCRRKLDITLNERDV 686
Qy 327 ESCSTENLE-LRKKEVLENTNRTLQLOKLQTL---VMGKV 365
Db 687 EDLQDKARLLREKVEFL-----RSLRQMKQKQVLSYQEVFGRL 725

RESULT 14
APL_CHICK STANDARD; PRT; 310 AA.
AC APL1CHICK
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Transcription factor AP-1 (Proto-oncogene c-jun).
GN JUN.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
[1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=94167107; PubMed=2577867;
RA Nishimura T., Vogt P.K.;
RT "The avian cellular homolog of the oncogene jun.";
RL Oncogene 3:659-663(1988).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS AND RECOGNIZE THE
CC -!- HANCER DNA SEQUENCE: TGA(C/G)TCA.
CC -!- SUBUNIT: INTERACTS WITH C-FOS TO FORM A DIMER.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. JUN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M57467; AAA48927.1; -.
DR HSSP; P05412; 1FOS.
DR TRANSFAC; T00134; -.
```

```
DR InterPro; IPR002112; Leuzip Jun.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; BZIP; 1.
DR SMART; PR00043; LEUZIPRJUN.
DR PRINTS; SM00338; BRUZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW Proto-oncogene; Transcription regulation; DNA-binding;
KW Nuclear protein.
FT DNA_BIND 236 255 BASIC MOTIF.
FT DOMAIN 259 287 LEUCINE-ZIPPER.
SQ SEQUENCE 310 AA; 33940 MW; 3EB36B31F96C994E CRC64;

Query Match
Best Local Similarity 5.9%; Score 156.5; DB 1; Length 310;
Matches 72; Conservative 56; Mismatches 112; Indels 101; Gaps 12;

Qy 74 APLQAEHSYSLSEPRTPQSPFTHAATSDSFNDEEV-----ESEKWL-----S 117
Db 27 AKVLQOSMTLNSDAASSLKPHLRKNADILTSVDGLLKLASPELERLIQSSNGLITT 86
Qy 118 TEFPSATTKKEPIEEOQ---PPGLVPSV-----TL-TITAISTPFKEEESPLDMNAGG 166
Db 87 TPTPTQFLCPKNVTDEQGFAGFVRALAEHLNQNTLPSTVSAAPQVSGMAPVSSMAGG 146
Qy 167 DSSCOTLPIKLEPHEVDQFNFSPKEAS-----VDQLHLPPTPPSSHSSDS 214
Db 147 GSFNTSL-----HSEPPVYANLSNFNPNALNSAPNANGMYAPQHINPQVQH---- 198
Qy 215 EGSLSNPRLHPFSLQAHSPVRAMPGRPSALSTSPLLTAPHKLQSGGLVLTEEEKRTL 274
Db 199 -----PRLQ--ALKEBPQVPEMPG-----ETPPL----- 221
Qy 275 VAGYPIPTKPLTKSEKALKIRKIKNKISAEQRRKKKEYMDSLEKKKEYSCSTENL 334
Db 222 -----SPIDMESQERIKAEKRMNRIRIAASKCRKLERIARLEEKVKTLKAQNS 271
Qy 335 ELRKKVEVLENTNRTLQLOKLQTLVMGKYSRTCKLAGTQ 375
Db 272 E-----LASTANMLRQVQALQKQKVMNVHNSGCOLMLTQ 305

RESULT 15
NFE2_HUMAN STANDARD; PRT; 373 AA.
AC Q16621; Q07720;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor NF-E2 45 kDa subunit (Nuclear factor, erythroid-
DE derived 2 45 kDa subunit) (P45 NF-E2) (Leucine zipper protein NF-E2).
GN NFE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94068604; PubMed=8248255;
RA Chan J.Y., Han X.L., Kan Y.W.;
RT "Isolation of cDNA encoding the human NF-E2 protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:11366-11370(1993).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93360994; PubMed=8355703;
RA Ney P.A., Andrews N.C., Jane S.M., Safer B., Purucker M.E.,
RA Weremowicz S., Goff S.C., Orkin S.H., Neinhuis A.W.;
RT "Purification of the human NF-E2 complex: cDNA cloning of the
RT hematopoietic cell-specific subunit and evidence for an associated
RT partner.";
RL Mol. Cell. Biol. 13:5604-5612(1993).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal liver;
```

Search completed: March 5, 2003, 06:19:23
Job time : 46 secs

```

RX MEDLINE=95241531; PubMed=7724591;
RA Pischedda C., Cocco S., Melis A., Marini M.G., Kan Y.W., Cao A.,
RT Moï P.;
RL "Isolation of a differentially regulated splicing isoform of human
NF-E2.";
CC Proc. Natl. Acad. Sci. U.S.A. 92:3511-3515(1995).
CC -|- FUNCTION: REQUIRED FOR ACTIVITY AT THE LOCUS CONTROL REGION (LCR)
CC UPSTREAM OF THE GLOBIN GENE COMPLEXES. REQUIRES P18 NF-E2 FOR
CC BINDING TO THE NF-E2 MOTIF. MAY PLAY A ROLE IN ALL ASPECTS OF
CC HEMOGLOBIN PRODUCTION; GLOBIN SYNTHESIS, HEME SYNTHESIS, AND THE
CC PROCUREMENT OF IRON.
CC -|- SUBUNIT: HETERODIMER OF P45 NF-E2 AND P18 NF-E2.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- TISSUE SPECIFICITY: EXPRESSED IN HEMATOPOIETIC CELLS AND ALSO IN
CC COLON AND TESTIS.
CC -|- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L24122; AAA16118.1; -
DR EMBL; L13974; AAA35612.1; -
DR EMBL; S77763; AAB34115.1; -
DR HSSP; P34707; 1SKN.
DR TRANSFAC; T01440; -
DR Genew; HGNC:7780; NFE2.
DR MIM; 601490; -
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; BZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
FT DOMAIN 55 62 POLY-PRO.
FT DOMAIN 77 82 POLY-PRO.
FT DNA_BIND 268 287 BASIC MOTIF.
FT DOMAIN 294 339 LEUCINE-ZIPPER.
FT CONFLICT 114 114 G -> A (IN REF. 2).
FT CONFLICT 214 214 A -> R (IN REF. 2).
FT CONFLICT 334 335 FQ -> LE (IN REF. 2).
SQ SEQUENCE 373 AA; 41472 MW; A9821170FB2ED67C CRC64;

Query Match 5.8%; Score 152.5; DB 1; Length 373;
Best Local Similarity 24.1%; Pred. No. 0.08;
Matches 78; Conservative 30; Mismatches 104; Indels 111; Gaps 12;

QY 118 TFPFSATIKKEPTTEQPP---GLVPSVLTITAI-----STPPEKESPLDMNA 164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 TELQGLNAPSESPFQAPAPYLGPPPTTYCFCSIHPSDGFPLPPPYSLPASTSHVPD 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 165 GGDSSCQTLPKIKLEPHVDQFLNFPKESVDQLHL-----PPTPPSSHSDS-- 214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 PPYSYGNMAIPVSK--PLSLGGLLS-----EPLQDPLALLDIGLPAGPKQEDPESDGL 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 215 -----EG-----SLSPNPRLH-----PFSLSQA 232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 SLNYSDAESLELEGTEAGRRRSEVEMYPVEYPSLMPNSLAHSNYTLPAETPLALEPS 208
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 233 HSPVRAMP--RGPSALSTSPLLTAPHKLQSGGPLVLTEEEKRTLVAEGYIPT----- 283
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 SGPVRAKPTARGEAG-----SRDERRAL-AMKIPPTDKIVNLP 246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 284 -----KLPLTKSEKALKIRIKKNSIAQESRRKKKKEYMDSLEKKVESCSTENL 334
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 "VDDFNELLARYPLTESQALVRDIRRGKGNKVAQAQNCRRKRLTIIVQLELESLTNERE 306
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 335 ELRKKVEVLENTNRTLQLOKL 357
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 RLLRARGEADRTLVEVMRQQLTEL 329
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

QY 206 PPSHSDSGSLSPNRLHPFSLQAHSPVRAMPGPSALSTPLTAPHKLOG-SGPL 264
DB 208 PSSHSGSDSGSQSP-SLPP-----SPVPRMARSTAISTPLTAPHKLOGTSGPL 260
QY 265 VLTEBEKRTLVAEGYPIPTKLPITKSEKALKIRKIKNKISAQESRRKKKEYMDSLEK 324
DB 261 LUTEBEKRTLIAEGYPIPTKLPITKAEKALKVRKIKNKISAQESRRKKKEYVECLEK 320
QY 325 KVESCSSTENLELRKKVELENTNRTLLQOLQOTLVMGKVSRTCKLAGTGTGCLMVVV 384
DB 321 KVFETFTSENELWKKVETLENARTLLQOLQOTLVNKRIPYKMAATQTGTCLMVAA 380
QY 385 LCFVAVFGSFFQGYGYPSPATKMALPSQHPLSEP--YTASVVRNLLIYEEHAPLEESS 442
DB 381 LCFVLVLGSLVPCLPFFSGSQTV--KEDPLAADGVYTASQMPSPSRLLFYDDGA----- 432
QY 443 SPASTGELGWDGRGSSLL 460
DB 433 -----GLWEDGRSTL 442

RESULT 2
Q96CP0
ID Q96CP0 PRELIMINARY; PRT; 519 AA.
AC Q96CP0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Similar to old astrocyte specifically induced substance.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC TISSUE=COLON;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014097; AAH14097.1; -.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; BZIP; 1.
DR PROSITE; PS00036; BZIP_BASIC; UNKNOWN 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 519 AA; 57031 MW; D1313PFOA02A6A8 CRC64;

```

Query Match 35.2%; Score 932.5; DB 4; Length 519;
 Best Local Similarity 50.0%; Pred. No. 2.1e-55;
 Matches 219; Conservative 52; Mismatches 118; Indels 49; Gaps 14;

```

DB 35 HFSLELDDEFSQNLGQLLSDPPLSEKSESEMEVEPSPPTAPLIQAHSYSLSEPRTPSP 94
DB 42 HFTENMEDFSNDLFFSSFFDDPVLDEKSLDLMELD--SPTPGIQAHSYSLSGDSAPQSP 99
QY 95 FTHAATSDFNDEBESEKWLSTFPFSATIKKEPITEEQP--PGLVPSVTLTITAITST- 151
DB 100 LVPVKMEDTTQD--AEHGAWALGHLKLSINVKQEQ-SPELPVDPPLAAPSAAAAAATTT 156
QY 152 -----PFKEESPLDMNAGDSSCQTLIPKILPEHEVDQFLNFSKPEASVDQLHLPPT 205
DB 157 PLGLGLSPLSR--LPHPQAPGE---MTQLPVKAEPLEVNQFLKVTPE---DLVQMPPT 207
QY 206 PPSHSDSGSLSPNRLHPFSLQAHSPVRAMPGPSALSTPLTAPHKLOG-SGPL 264
DB 208 PSSHSGSDSGSQSP-SLPP-----SSVPRMARSTAISTPLTAPHKLOGTSGPL 260
QY 265 VLTEBEKRTLVAEGYPIPTKLPITKSEKALKIRKIKNKISAQESRRKKKEYMDSLEK 324
DB 261 LUTEBEKRTLIAEGYPIPTKLPITKAEKALKVRKIKNKISAQESRRKKKEYVECLEK 320
QY 325 KVESCSSTENLELRKKVELENTNRTLLQOLQOTLVMGKVSRTCKLAGTGTGCLMVVV 384
DB 321 KVFETFTSENELWKKVETLENARTLLQOLQOTLVNKRIPYKMAATQTGTCLMVAA 380
QY 385 LCPAFAFGSFFQGYGYPSPATKMALPSQHPLSEP--YTASVVRNLLIYEEHAPLEESS 442

```

```

DB 381 LCFVLVLGSLVPCLPFFSGSQTV--KEDPLAADGVYTASQMPSPSRLLFYDDGA----- 432
QY 443 SPASTGELGWDGRGSSLL 460
DB 433 -----GLWEDGRSTL 442

RESULT 3
Q92I25
ID Q92I25 PRELIMINARY; PRT; 507 AA.
AC Q92I25;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE OASIS protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ICR;
RC MEDLINE=99280577; PubMed=10350641;
RA Homma Y., Kanazawa K., Mori T., Tanno Y., Tojo M., Kiyosawa H.,
RA Takeda J., Nikaide T., Tsukamoto T., Yokoya S., Wanaka A.;
RT "Identification of a novel gene, OASIS, which encodes for a putative
RT CREB/ATF family transcription factor in the long-term cultured
RT astrocytes and gliotic tissue."
RL Brain Res. Mol. Brain Res. 69:93-103(1999).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY.
DR EMBL; AB017614; BAA75670.1; -.
DR MGD; MGI:1347062; Oas1.
DR InterPro; IPR001630; Leuzip CREB.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; BZIP; 1.
DR PRINTS; PR00041; LEUZIPPCCREB.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 507 AA; 55428 MW; AAE2F859DAA1D978 CRC64;

```

Query Match 35.0%; Score 928; DB 11; Length 507;
 Best Local Similarity 47.9%; Pred. No. 4.1e-55;
 Matches 221; Conservative 50; Mismatches 138; Indels 52; Gaps 13;

```

QY 35 HFSLELDDEFSQNLGQLLSDPPLSEKSESEMEVEPSPPTAPLIQAHSYSLSEPRTPSP 94
DB 42 HFTENMEDFSNDLFFSSFFDDPVLDEKSLDLMELD--SPAPGIQAHSYSLSGDSAPQSP 99
QY 95 FTHAATSDFNDEBESEKWLSTFPFSATIKKEPITEEQP--PGLVPSVTLTITAITSTP 152
DB 100 LVPVKMEDTTQD--VEHGAWALGHLKLSINVKQEQ-SPELPVDPPLAASSAAAAAATP 156
QY 153 -----PFKEESPLDMNAGDSSCQTLIPKILPEHEVDQFLNFSKPEASVDQLHLPPT 207
DB 157 PLGLGLSPMLRLPHPQAPGE---MTQLPVKAEPPEMSQFLKVTPE---DLVQMPPT 209
QY 208 PSSHSDSGSLSPNRLHPFSLQAHSPVRAMPGPSALSTPLTAPHKLOG-SGPLVL 266
DB 210 PSSHSGSDSGSQSP-SLPP-----SSVPRMARSTAISTPLTAPHKLOGTSGPL 262
QY 267 VLTEBEKRTLVAEGYPIPTKLPITKSEKALKIRKIKNKISAQESRRKKKEYMDSLEK 326
DB 263 LUTEBEKRTLIAEGYPIPTKLPITKAEKALKVRKIKNKISAQESRRKKKEYVECLEK 322
QY 327 KVESCSSTENLELRKKVELENTNRTLLQOLQOTLVMGKVSRTCKLAGTGTGCLMVVV 386
DB 323 ETVTSENELWKKVETLENARTLLQOLQOTLVNKRIPYKMAATQTGTCLMVAA 382
QY 387 PAVAFGSPFQGYGYPSPATKMALPSQHPLSEP--YTASVVRNLLIYEEHAPLEESSPAS 446

```

```
Db 383 FVLVLGSLVPCLPAPFSGSGMTVKEDPIAADSYYAASQMPRSRLFFYDDGA----- 432
Qy 447 TGEICGWD--RGSSII-----LRASGLEALPEVDLPH 476
Db 433 ----GSWEDGRGALLFVPEPEGWELKPGGPAEQRPQDLRLH 469

RESULT 4
Q91W70 PRELIMINARY; PRT; 520 AA.
ID Q91W70
AC Q91W70
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Old ascocyte specifically induced substance.
GN OASIS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary Gland;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016447; AAH16447.1; -.
DR MGD; MGI:1347062; Oas1.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; UNKNOWN 1.
SQ SEQUENCE 520 AA; 57087 MW; DC2652AA1583A9F2 CRC64;

Query Match 35.0%; Score 928; DB 11; Length 520;
Best Local Similarity 47.9%; Pred. No. 4.3e-55;
Matches 221; Conservative 50; Mismatches 138; Indels 52; Gaps 13;

Qy 35 HFESELLDFSONVLGQLSDPFLSKSMEVEPSPAPLQAEHSYSLSEEPRTQSP 94
Db 42 HFVENMEDFSNDLFSFDDPVLDEKSLDLDWELD--SPAPGQAEHSYSLSGDSAQSP 99
Qy 95 FTHAATSDSFNDEEVESEKWLSTFPFGATIKKEPITEBP--PGLVPSVTLTITAIPT 152
Db 100 LVPVKMETTDQ--VEHGAWALGNKLCISIMVKQEQ-SPELPVDPLAASAMAAAAMATP 156
Qy 153 -----FEKEESPLDMNAGDSSCQTLIPKIKLEPHEVDQFLNFSPEASVDQLHLPPTPP 207
Db 157 PLLGLSPMRPLPIPHQAPGE---MTQLPVIKAEPPMSQFLKVTPE---DLVQMPPTPP 209
Qy 208 SSHSDSGSLSPNRLHPFLSLSQAHSPVRAMPGPSALSTSPLLTAPHKLOG-SGPLVL 266
Db 210 SSHGSDSGSQSPR-SLPP-----SSPVRPMARSGTAISTSPLLTAPHKLOGSGPLLL 262
Qy 267 TEEKRTIAEGYPIPTKLPTKSEKALKKTRKIKKISAQESRRKKKEYMDSLEKKV 326
Db 263 TEEKRTIAEGYPIPTKLPLTKABEALKRVRRKIKKISAQESRRKKKEYVECLEKKV 322
Qy 327 ESCSTENLELRKRVLENTNRTLQQLQKQLTLMVKGVSRCTKLAGTGTGCLMVMVLC 386
Db 323 EYTTSENELMKVKVETLETANRTLQQLQKQLTMTSKISRYKKAATGTGCLMVAALC 382
Qy 387 FAVAFSGFQGGYVPSAKMALPSQHLSEPYTASVVRNLLIYEHPLEESSSPAS 446
Db 383 FVLVLGSLVPCLPAPFSGSGMTVKEDPIAADSYYAASQMPRSRLFFYDDGA----- 432

Qy 447 TGEICGWD--RGSSII-----LRASGLEALPEVDLPH 476
Db 433 ----GSWEDGRGALLFVPEPEGWELKPGGPAEQRPQDLRLH 469

RESULT 5
Q24282 PRELIMINARY; PRT; 516 AA.
ID Q24282
```

```
AC Q24282;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE CRE-binding transcriptional activator (CREB-A).
GN CREBA OR CREB-A OR CG7450.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92375081; PubMed=1508208;
RA Smolik S.W., Rose T.M., Goodman R.H.;
RT "A cyclic AMP-responsive element-binding transcriptional activator in
RT Drosophila melanogaster, dCREB-A, is a member of the leucine zipper
RT family.";
RL Mol. Cell. Biol. 12:4123-4131(1992).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE bZIP FAMILY.
DR EMBL; M87038; AAA28427.1; -.
DR FlyBase; FBgn0004396; CrebA.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRL2; 1.
DR PROSITE; PS00036; bZIP_BASIC; 1.
DR DNA-binding; Nuclear protein.
SQ SEQUENCE 516 AA; 56281 MW; 15BBF2108BE53983 CRC64;

Query Match 16.5%; Score 436; DB 5; Length 516;
Best Local Similarity 33.7%; Pred. No. 1.1e-21;
Matches 142; Conservative 54; Mismatches 117; Indels 108; Gaps 16;

Qy 26 ETEALMYH---TFSELDFSONVLGQLSDPFLSKSMEVEPSPAPLQAEHS 82
Db 111 DPKCLTFHVPPHATP-----ISRLSSNPALN-----TSVADLTFSSGL 149
Qy 83 YSL-SEEPRTQSPFTHAATSD-----SPNDEEVESEKWL-----STEPS 122
Db 150 QSLQAHQPHHGSGSHVVVANLEHFQPLQHLVDNDCSSSVSLDRGSMSPDICSDIE 209
Qy 123 ATKKEPITEBPGLVPS-----VTLTITAIPT--FEKEESPLDMNAGDSSCQ 171
Db 210 SAIKDEPMSDSCPASPTSQASSQHQLSLNLAHQSEMLFEPKHCGLLTASSNSN-N 268
Qy 172 TLIPK-----IKLEPHEVDQFLNFSPEASVDQLHLPPTPSS--HSSDSE 215
Db 269 SLIKSQQQQQILGQDNLMMKMEIKSEKQSTNSSDSKSHAGYGIPITPPSSLSFSD 328
Qy 216 GLSLSPNRLHPF-----SLQAH-----SPVR-----AMRPGPSALST- 248
Db 329 GNLSPHLFAPLSNPATVSIIVANPAGGESSVRVSRTAASITRSSSGSASGSSSTSV 388
Qy 249 -----SPLLTAPHKLOGGGLVLTBEERKTLVAEGYPIPTKLPTKSEBALKKTRR 300
Db 389 TTTROPITPLISSQPK-GSTGTLTTEERKTLIAEGYPIPKLPTKAEBSLKIRR 447
Qy 301 KIKNKISAQESRRKKKEYMDSLEKVESCTENLELRKRVLENTNRTLQQLQKQLT 360
Db 448 KIKNKISAQESRRKKKEYMDQLRRVEILVTENHDYKKRLEGBEETNANLSQLHQL 507
Qy 361 V 361
Db 508 V 508

RESULT 6
Q9VUQ4 PRELIMINARY; PRT; 518 AA.
ID Q9VUQ4
AC Q9VUQ4;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
```


01-MAR-2002 (TREMREL. 20, Last annotation update)
CRSEA protein.
CRSEA OR C67450.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_taxId=7227;
RN
RN
SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.A., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hooten D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster";
Science 287:2185-2195(2000).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY.
DR EMBL; AB003530; AAF49621.1; -.
DR FlyBase; FBgn004396; Creba.
DR InterPro; IPR001630; Leuzip_CREB.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; bzip; 1.
DR PRINTS; PRO0041; LEUZIPPRCREB.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 518 AA; 56527 MW; 9461C27F2AB7E26 CRC64;

Query Match 16.4%; Score 435; DB 5; Length 518;
Best Local Similarity 34.8%; Pred. No. 1.3e-21;
Matches 138; Conservative 50; Mismatches 112; Indels 96; Gaps 14;

QY 48 LQQLSDPFLSEKSEMEVPSPTSPAPLQAEHSYL-SEEPRTQSPFFTHAATSD---- 102
DB 129 ISRLSNPALN-----TSVADLTRSSGLQSLQAHQHGSGSHVVVANLEHF 176
QY 103 ----SFNDEEVESEKWL-----STEPSATIKKEPITEEQPPGLVPS----- 141

Db 177 QLPQHLYDNCSSVSLRDGSMSPDICSDIEIDESAIKDBPMSPDSSCPASPTSQASS 236
QY 142 ---VTLTTAISTP--FEKESPLDMNAGDSSCOTLIPK-----IKLEP 181
Db 237 QHQLSLNLAHQSEMLFPFKHGLLTASSNSN-NSLTKSQORQOQIIGQDNLMAKMEI 295
QY 182 HEVDQFLNFPKPEASVDQHLHPPTPPSS-HSDDSGSLSPNPLHPF-----SLSQAH- 233
Db 296 KSEKOSTNSNDKSHAGYGLPTPTPSSLPDSDSGNLSPEHLFAPLSNATVSVANP 355
QY 234 ----SPVR-----AMPROPSALST-----SPLLTAPHKLGSGPLV 265
Db 356 AGCESSVRVSRTAASITRSSSGSASGSSTSTVTTTRQPIHTPLISSQPK-GSTGTLL 414
QY 266 LTEREKRTLVAGYDIPTKLPTKSEKALKIRKKNKISAQESRRKKKYMDSLEKK 325
Db 415 LTEREKRTLVAGYDIPQKPLTKAEKSLKIRKKNKISAQESRRKKKYMDSLEKK 474
QY 326 VESCSTENLELRKKVELENTNRTLLQQLQQLTV 361
Db 475 VEILTENHDYKKRLEGTNANLLSQLKQALV 510

RESULT 7
Q8TEYS PRELIMINARY; PRT; 395 AA.
ID Q8TEYS
AC Q8TEYS
DT 01-JUN-2002 (TREMREL. 21, Created)
DT 01-JUN-2002 (TREMREL. 21, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE Androgen-induced basic leucine zipper.
GN AIBZIP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21818580; PubMed=11830526;
RA Qi H., Fillion C., Labrie Y., Grenier J., Fournier A., Berger L.,
RA El-Alfy M., Labrie C.;
RT "AIBZIP, a Novel bZIP Gene Located on Chromosome 1q21.3 That Is Highly
RT Expressed in Prostate Tumors and of Which the Expression Is Up-
RT Regulated by Androgens in LNCaP Human Prostate Cancer Cells";
RL Cancer Res. 62:721-733(2002).
DR EMBL; AF394167; AAL76113.1; -.
SQ SEQUENCE 395 AA; 43432 MW; F3DDF288421AB5F2 CRC64;

Query Match 14.9%; Score 394; DB 4; Length 395;
Best Local Similarity 34.0%; Pred. No. 5.8e-19;
Matches 115; Conservative 49; Mismatches 108; Indels 66; Gaps 12;

QY 135 PGLVPSVTLTITATSTPFKEBESPLDMNAGDSSC--QTLIP-----KIKLEPEV--DQ 186
Db 33 PPPEVPVTRL-----QEQLQGWKSGDGRGCLQSEDEFLKLFIDNEVYVSE 82
QY 187 FLNFPKPEASVDQHLHPPTPPSSHSDS-----EGSLSPNPLHPFSL 230
Db 83 ASFGSDSGISEDPCF-PDGPAPRATSPMLYEVVYVYAGALERMQGTGPNVGLISQLD 141
QY 231 QAHSVPVRAMPGRGPSALSTSPLLTAPHKLGSG-----PLVTEBEKRTLVA 276
Db 142 Q-WSPAFMVP-DSCWSELFPDHAHLPRAGTVAPVCTTLPCQTLFTDEKRLGQ 199
QY 277 EGYPIPTKPLTKSEKALKIRKKNKISAQESRRKKKYMDSLEKKYVESCSTENLE 336
Db 200 EGVSLPSHLPLTKAEERVLKVRKKNKQSAQDSRRKKKEYIDGLSESVACSAQOEL 259
QY 337 RKKEVLENTNRTLLQQLQQLQQLTVMGKVBRTCKLAGTGTCTCLMVVVLFAVAFGFFQ 396
Db 260 QKKVQLERHNI SLVAQLQLOTL-----IAQTSNKA-AQTSTCVLILLFSLAILPSFS 314
QY 397 GYGPYSATKWLPSQHPLEPVTASVVRNLLIYEE 434


```
Db 315 PQSRPEAG-----SEDYQPHGVTSRNILTHKD 342
RESULT 8
Q96TB9 PRELIMINARY; PRT; 461 AA.
AC Q96TB9
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CREB/ATF family transcription factor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21267166; PubMed=11353085;
Watanabe S., Tanigami A., Sugano S.;
Omori Y., Imai J., Watanabe M., Komatsu T., Suzuki Y., Kataoka K.,
"CREB-H: a novel mammalian transcription factor belonging to the
CREB/ATF family and functioning via the box-B element with a liver-
specific expression.";
Nucleic Acids Res. 29:2154-2162(2001).
DR EMBL; AB050902; BAB47242.1; -.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; UNKNOWN 1.
SQ SEQUENCE 461 AA; 49077 MW; C40F5AA626661F80 CRC64;

Query Match 14.5%; Score 384.5; DB 4; Length 461;
Best Local Similarity 28.4%; Pred. No. 3.1e-18;
Matches 148; Conservative 70; Mismatches 175; Indels 129; Gaps 19;

Qy 1 MEVLESGEQSVLOWDRKLSLSEPCETALMYHVFSELLDEF--SONVLGQLSDPFLSE 59
Db 17 MDPIDSFELLDLLFDQ-----DGLRNVELAGWILAREEQKVLNLSDSDEFL- 66
Qy 60 KSESMVEPSPPTAPLIQAHSYLSSEPRTPQSFTHAATSDSFNDEVESEKWLST 119
Db 67 ILGSGDSLPS-----SPLWSPGSDSGISEDLPD----- 96
Qy 120 FPSATIKKEPTTEOPGLVPSVTLTITATSTPFEKESP-LDMNAGDSSCQTLIPKIK 178
Db 97 -----PQDTPPRSFGA---TSPAGCHPAQPGKPCLSYHPG--NSCSTTTTGPV 140
Qy 179 LEPHEVDQFLN---PSP-----KEASVDOLHLP---TPSSHSDSGSLSPNRLHP 226
Db 141 IQVPEASVTIDLEWSPGGRICAEKPADPVDLSRPNLTVDLLSGSGDL----- 192
Qy 227 FSLQAHSFVRAMPGRPSALSTSPLLTAPHKLGSG---PLVLTTEEKRTLVAEGYPT 283
Db 193 ----QQRH-----LGASYLLR-----PGAGHCQELVLTDEKLLAKEGITLPT 232
Qy 284 KLPLTKSEKALKIRRKIKNKISAOESRRKKKYMDSLEKKVSCSTENLELRKKVYL 343
Db 233 QLPLTKYEERVLKIRRKIRNKQSAQESRKKKKYIDGLTRMSACTAQNOELQKVLHL 292
Qy 344 ENTNRTLQQLQKLTLMVGKVSRTCKLAGTGTGTCMLVVLFCFAVAFSGPFGYGP 403
Db 293 EKQNLSLLEQLKQLQAVVQSTSKS-----AQTGTCVAVLLSFLAILPLPSIFGP 344
Qy 404 ATKNALPSQHPLEPYTASVVRNLLIYEHAPLEES-----SSPASTGELG 451
Db 345 -NKTESPGDFAPVRFSTLHNDASRVAADAVFQSEAPGRPEADTTREESPGSGADW 403
Qy 452 GWRDGSLLRASSGLEALPEVDLPFLISNET-SLEKSVLLE 492
Db 404 GFQDTANLTNSTE-----ELDNATLVLRNATEGIGQVALLD 439

RESULT 9
```

```
Q91XE9 PRELIMINARY; PRT; 479 AA.
AC Q91XE9
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Similar to CREB/ATF family transcription factor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=LIVER;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010786; AAH10786.1; -.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; UNKNOWN 1.
SQ SEQUENCE 479 AA; 52145 MW; 948A06FE79DB3779 CRC64;

Query Match 14.3%; Score 378.5; DB 11; Length 479;
Best Local Similarity 29.0%; Pred. No. 8.3e-18;
Matches 157; Conservative 62; Mismatches 153; Indels 159; Gaps 24;

Qy 1 MEVLESGEQSVLOWDRKLSLSEPCETALMYHVFSE--LLDEFSONVLGQLSDPFLS 58
Db 17 MAPLDSMEVLDLLFDQ-----DGLRNVELAGWILAREEQKVLNLSDSDEFL- 65
Qy 59 KSESMVEPSPPTAPLIQAHSYLSSEPRTPQSFTHAATSDSFNDEVESEKWLST 118
Db 66 ----NCILPGDSDPS-----SPLWSPADSDSGISEDLPD----- 97
Qy 119 EFPSATIKK---EPI-----TEEOPGLVPSVTLTITATSTP-FEKEES---PLDMN 163
Db 98 --PQDTPPRSCTEPANTVARCTREQGKPCPSYLPSTPCPEPPTQVQESSVAIDLDM- 154
Qy 164 AGDSSSCQTLIPKIKLEPHEVDQFLNFPKESASVDQLHLPTTPSSHSDSGSLSPNRP 223
Db 155 ----WSTDTLYPE---EPAGSPSRFNLTVKELL-----SGSGDL----- 188
Qy 224 LHPFSLQAHSFVRAMPGRPSALSTSPLLTAPHKLGSG---PLVLTTEEKRTLVAEGY 280
Db 189 -----QQHSLAASQLLGP-----GSGHCQELVLTDEKLLAKEGVT 225
Qy 281 IPTKLPLTKSEKALKIRRKIKNKISAOESRRKKKYMDSLEKKVSCSTENLELRKKV 340
Db 226 LPTQLPLTKYEERVLKIRRKIRNKQSAQESRKKKKYIDGLTRMSACTAQNOELQKRV 285
Qy 341 EVLENTRTLQQLQKLTLMVGKVSRTCKLAGTGTGTCMLVVLFCFAVAFSGPFGYGP 400
Db 286 LHLEKQNLSLLEQLKHLQALVVQSTSKP-----AHAGTCIAVLLLSFAL----- 329
Qy 401 YPSATKNALPSQHPLESE-----PYTASVVRNLLIYEHAP-----LESSSPAS 446
Db 330 -----IILPSISPNKVDSPGDFVPRVFSRTL---HNHAASRVAPDVTPTGSEVGP 380
Qy 447 TGEIIGWDRGSSLLRASSGLEA-----LP-----EVDLPFLISNET-SLEKSVLL 491
Db 381 WPDVGTGTHKGPS-----SGGSADWGNFLEIPMLDNLTTELDNSTVLANSTEDLGRATLL 436
Qy 492 E 492
Db 437 D 437

RESULT 10
Q9D2A5 PRELIMINARY; PRT; 370 AA.
AC Q9D2A5
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
```

```

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 5330432F22R1k protein (RIKEN CDNA 5330432F22 gene).
GN 5330432F22R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PITUITARY GLAND;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guetincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerte P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SALIVARY GLAND;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK019928; BAB31922.1; -
DR EMBL; BC022605; AAB22605.1; -
DR MGD; MGI:1925534; 5330432F22R1k.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; BZIP; 1.
DR SMART; SM00338; BRLZ; 1.
SQ SEQUENCE 370 AA; 41024 MW; F64D7F2F14547B4F CRC64;

Query Match 13.4%; Score 355; DB 11; Length 370;
Best Local Similarity 30.9%; Pred. No. 2.4e-16;
Matches 120; Conservative 44; Mismatches 110; Indels 114; Gaps 15;

160 LDMNAGDSQCQLIPKIK-LEPHEVDQFLNFPKESVDQLHLPTPPSSHSSDSEGLS 218
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
22 LEIGFNGPAS---KVPVTRGLQKSEPDFLNLFP---IDPNMHCSETSPGRDGVSEDPG 75
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

219 SP-----NPLRH-----PFSL----- 229
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

76 SPAQAASSPALVEVYVDSGLTQGTQREAGPTGLISIQIDQWTPALMVPDCTVSGLPS 135
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

230 -SQAHSPVRAMPGRPSALSTSPLLTAPHKLGSGPLVLTBEKRTLVAGGYPTKLPPLT 288
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

136 DSHRHILPRVSTRAPPAAMP---SCQHH-----LFLTDEKQLLAQEGITLPSHLPLT 187
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

289 KSEKALKTKIRKKNKISAQESRRKKKEYMDSLEKVCSCSTENLELRKKEVLENTNR 348
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

188 KABERILKTKIRKKNKQAQSDRRKKKEYLDGLSEKVAACSEQNKQKLRKVQELERQNI 247
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

349 TLQLQQLKLTVMGKVRCTKLAGTGTCTCLMVVLCFVAFVFGSPFGYGPYPATKMA 408
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

248 FLMEQVRLQKLTAQTSR-----AAQTSTC--VLILLFSLAL-----II 285
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

409 LPSCHPLS-----EPTASVVRNLLIYEEHAPLEESSSPASTGELGWDGSGSLR 461
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

286 LPSFSPFGQGEARPEYQLHGVISRNILTHEN--VTENLESVP----- 327
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

462 ASSGLEALPEVDLPHFLISNETSLEKSV 489
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 328 LKSKLEELPEA--PTTNGSTKTLKMRV 353
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
Q9DAE0
ID Q9DAE0 PRELIMINARY; PRT; 315 AA.
AC Q9DAE0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 1700012K17R1k protein (Attaches to Cre).
GN 1700012K17R1K OR ATCE1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guscincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerte P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Stelzer G., Don J.;
RT "Accl - a novel protein that contains a CRE binding domain.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK005918; BAB24315.1; -
DR EMBL; AF287260; AAK73568.1; -
DR MGD; MGI:1916603; 1700012K17R1k.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; BZIP; 1.
DR SMART; SM00338; BRLZ; 1.
SQ SEQUENCE 315 AA; 35020 MW; BCE2B0EBBA240DD16 CRC64;

Query Match 13.0%; Score 344.5; DB 11; Length 315;
Best Local Similarity 33.2%; Pred. No. 1e-15;
Matches 107; Conservative 41; Mismatches 93; Indels 81; Gaps 11;

175 PKIKLEPHEVDQFLNFPKESVDQLHLPTPPSSHSSDSEGLSPNRLHPFSLSOAHS 234
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 PTFGLLSIQIDQ---WTPALMVPDCTVSGLPSDSHR-----HILPRVSTRAPA 96
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

235 PVAMPGRGPSALSTSPLLTAPHKLGSGPLVLTBEKRTLVAGGYPTKLPKTSSEKA 294
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

97 PPAAMP-----SCQHH-----LFLTDEKQLLAQEGITLPSHLPTKAEERI 138
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

295 LKIRKKNKISAQESRRKKKEYMDSLEKVCSCSTENLELRKKEVLENTNRILQQL 354
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

139 LKIRKKNKQAQSDRRKKKEYLDGLSEKVAACSEQNKQKLRKVQELERQIFLMEQV 198
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

355 QKLQTLVMGKVRCTKLAGTGTCTCLMVVLCFVAFVFGSPFGYGPYPATKVALPSQHP 414
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

199 RLQKLTAQTSR-----AAQTSTC--VLILLFSLAL-----IILPSFSP 236
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
QY 415 LS-----EPYASVVRNLLIYEEHAPLESSSPASTGELGWDGSSLLRASSGLE 467
DQ 237 FQGOSEARPEYQLHGVISRNILTHEN--VTENLESPV-----LKSLE 278
QY 468 ALPEVDLPFLHISNETSLEKSV 489
DQ 279 ELPEA--PTTNGSTKHLKMRV 298

RESULT 12
Q8SQ19 PRELIMINARY; PRT; 365 AA.
AC Q8SQ19;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Luman.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Budihal P.C., Misra V.;
RL Submitted (May-2001), to the EMBL/GenBank/DBJ databases.
DR EMBL; AF387035; AAL84006.1;
SQ SEQUENCE 365 AA; 40556 MW; 4351DCBA066F9822 CRC64;

Query Match 12.7%; Score 337; DB 6; Length 365;
Best Local Similarity 33.7%; Pred. No. 3.9e-15;
Matches 101; Conservative 44; Mismatches 103; Indels 52; Gaps 9;

QY 198 DQLHLPPTPPS--SHSSDSEGLSPNRLHPFSLQAHSPV-----RAMPRGPSA 245
DQ 52 DFLSCLPSPPAVLNFSNSDPLVQHD--HTYLSQEHVSIDLNESYKERA----- 102
QY 246 LSTGPLLTAHKLQGGPLVITEEKRTLVAAGYPIPTKLPTKSEKALKIRKIRK 305
DQ 103 -QMTPLRVEEPADQIARLLITEEKRLLEKGLTLPGLPLTKMEEQVLKRVRRKIRK 161
QY 306 ISAOESREKKKEYMDSLEKVESCTENLELRKKVEVLNTNRTLLQLOKLOTLMVK 365
DQ 162 KSAOESRRKKKYVGGLESRLKYTAQNLQONKVLLEEQNLSDLLDQLRLQAMVIQTA 221
QY 366 SRTCKLAGTGTCTCLMVVLCFAVAFGFFGYGYPATKALPSPHPLSEPYTASVVR 425
DQ 222 NK-----ASSSTCVLLVLFSCILLVP-----AMYSDTRGSLPAEHR-----VL 262
QY 426 SRNLLIYEHAPLESSSPASTGELGWDGSSLLRASSGL-----EALPEVDLP 475
DQ 263 SRQLRALPSDEP-PQLPEPALQSEVPKDSLNPQLQASNSCCLFLHMPQAPRAEPPLQLP 321

RESULT 13
O14671 PRELIMINARY; PRT; 371 AA.
AC O14671;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Luman.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu R., Yang P., O'Hare P., Misra V.;
RL MEDLINE=97415590; PubMed=9271389;
RT "Luman, a new member of the CREB/ATF family, binds to herpes simplex
virus VP16-associated host cellular factor.";
```

```
RL Mol. Cell. Biol. 17:5117-5126 (1997).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY.
DR EMBL; AF009368; AAB69652.1;
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 371 AA; 41319 MW; 22042E496B92535A CRC64;

Query Match 12.5%; Score 330.5; DB 4; Length 371;
Best Local Similarity 30.7%; Pred. No. 1.1e-14;
Matches 116; Conservative 56; Mismatches 123; Indels 83; Gaps 15;

QY 150 STPEKEESPLDMAGDSSCQTLPKIKLEPHEVDQFLNFSPEKASVDQLHLP-PTPPS 208
DQ 25 TAPDEAVRAPLDW-----ALP-----LSEVPDWEVDLLCSLLSPPA 62
QY 209 SHSSDSEGLSPNRL--HPFSLQAHSPVAMPRGPSALSTSPLLTAPH-----KLGSG 262
DQ 63 SLNLS-----SSNPLVHHHTYSLPRETVSMDLSESCREKGTQMTPOHMEELAEQIEA 118
QY 263 PLVLTTEEKRTLVAEGYPIPTKLPTKSEKALKIRKIRKISAOESRRKKKEYMDSL 322
DQ 119 RLVTDEEKSILLEKGLILPETLPKTEEQILKRVRRKIRKISAOESRRKKKYVGG 178
QY 323 EKKVESCTENLELRKKVEVLNTNRTLLQLOKLOTLMVKGVSRCTCKLAGTGTCTCLMV 382
DQ 179 ESRVLKYTAQNMELQNKVLLEEQNLSDLLDQLRLQAMVIEISNKT-----SSS 233
QY 383 VVLCFAVAFGFFGYGYPATKALPSPH-----PLSEPYTASVVRNLLIY 432
DQ 234 LLVSFCILLVP-----AMYSDTRGSLPAEHLVSRQLRALPSEDPYQLEL----- 279
QY 433 EEHAPLESSSPA-STGELGWDGSS--LLRASSGLEAL-----PEVDLPHELI 479
DQ 280 ----PALQSEVPKDSLTHQ----WLDGSDCVLQAPNGTSCLLHYMFOAPSAEPPEPDL 332
QY 480 SNETSLEKSVLLELQOHL 497
DQ 333 SSE-PLCRGPILPQANL 349

RESULT 14
Q9H2W3 PRELIMINARY; PRT; 371 AA.
AC Q9H2W3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CAMP responsive element binding protein 3.
GN CREB3
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ben-Yosef T., Francomano C.A.;
RT "Complete nucleotide sequence and genomic structure of the human CAMP
responsive element binding protein 3 (Luman) gene (CREB3).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY.
DR EMBL; AF211847; AAG43527.1;
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 371 AA; 41451 MW; 243C4C041B3E1125 CRC64;
```

Query Match		12.4%;	Score 329.5;	DB 4;	Length 371;
Best Local Similarity		29.9%;	Pred. No. 1.3e-14;		
Matches 117;		Conservative 55;	Mismatches 122;	Indels 97;	Gaps 16;
QY	160	LDNAGGSSCQTLIPKIKLEPHEVDQFLNFPKASVDQLHLP---	PTPPSHSSDSE-	215	
DB	3	LELDAGDQDLIAFLI-----ESEDLTAPDEA----	VRAPLDWA	PLSEVPSDWEV	50
QY	216	-----GSLSPNRLHPFSLQA-----HSPVRAMR-----	GPSALSTSPLLTAPH--	256	
DB	51	DDLCSLLSPASLANLSSNCLVHHDTYSLPRETVSMDLSESCREKGTQMTPOHME	110		
QY	257	--KLGSGPLVTEBEKRTLVAGYPIPTKLPTKSEKALKIRKIKNKISAOESRRK	314		
DB	111	ELAEQEIARLVLTDEKSLLEKGLIPETLPTKTEEQILKVRKIRNKSAQESRRK	170		
QY	315	KKEYMDSLEKVESCTENLELRKKEVLENTNRTLLQQLQKLTVMGKVSRTCKLAGT	374		
DB	171	KVYVGLSERVLKYTAQNMELQNKVQLLEEQNLSDQLRKLAQMWIEISNKT-----S	225		
QY	375	QTGCLMVVLCFAVAFSGFFQGYGYPYSATKMALPSQH-----	PLSEPYTASVV	424	
DB	226	SSSTCILVLLVFCLLLP-----AMVSSDTRGSLPAEHGVLRSQRLALPSEDPYQLEL-	279		
QY	425	RSRNLLIYEHAPELESSSPA-STGELGWDGSSLLRASSG-----	LEALPEVDLPHEL	478	
DB	280	-----PALQSEVPKDSHQ-----WLDGSDCVLQAGNTSCLLHYMPQAP-----	319		
QY	479	IGNETSLE-----KSVLLELQOHL	497		
DB	320	-SAEPPLEWFPDLPSEPLCRGPILPQANL	349		

RESULT 15
O14919
ID O14919 PRELIMINARY; PRT; 371 AA.
AC O14919;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Basic leucine zipper protein LZIP (CAMP responsive element binding protein 3) (HUMAN).
DE LZIP OR CREB3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
SEQUENCE FROM N.A.
Freiman R.N., Herr W.;
Genes Dev. 11:0-0(1997).
RL [2]
RP SEQUENCE FROM N.A.
RA Ben-Yosef T., Francomano C.A.;
RT "Complete nucleotide sequence and genomic structure of the human CAMP responsive element binding protein 3 (luman) gene (CREB3).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY.
DR EMBL; AF029674; AAB84166.1; -.
DR EMBL; AF211848; AAC43528.1; -.
DR EMBL; BC010158; AAH10158.1; -.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 371 AA; 41379 MW; 82152E496B924BEC CRC64;

Query Match		12.3%;	Score 326.5;	DB 4;	Length 371;
Best Local Similarity		30.0%;	Pred. No. 2.1e-14;		
Matches 115;		Conservative 54;	Mismatches 121;	Indels 93;	Gaps 15;
QY	150	STPFKEKESPLDMNAGSSCQTLIPKIKLEPHEVDQFLNFPKASVDQLHLP-PTPPS	208		
DB	25	TAPDEAVRAPLDM-----ALP-----	LSEVPSDWEVDLLCSLLSPPA	62	
QY	209	SHSSDSEGLSPNRL--HPFSLQAHSVPVAMPRGPSALSTSPLLTAPH-----	KLGSG	262	
DB	63	SINILS-----SSNPCLVHHDTYSLPRETVSMDLSESCREKGTQMTPOHMEELAEQIA	118		
QY	263	PLVLTTEEKRTLVAGYPIPTKLPTKSEKALKIRKIKNKISAOESRRKKEYMDSL	322		
DB	119	RLVLTDEKSLLEKGLIPETLPTKTEEQILKVRKIRNKSAQESRRKKEYVVGGL	178		
QY	323	EKKVESCTENLELRKKEVLENTNRTLLQQLQKLTVMGKVSRTCKLAGTQTCMLV	382		
DB	179	ESRVLKYTAQNMELQNKVQLLEEQNLSDQLRKLAQMWIEISNKT-----SSSSTCILV	233		
QY	383	VVLCFAVAFSGFFQGYGYPYSATKMALPSQH-----	PLSEPYTASVVRNLLIY	432	
DB	234	LLVSPCLLIV-----AMVSSDTRGSLPAEHGVLRSQRLALPSEDPYQLEL-	279		
QY	433	EEHAPLESSPA-STGELGWDGSSLLRASSG-----	LEALPEVDLPHELISNETSLE	486	
DB	280	-----PALQSEVPKDSHQ-----WLDGSDCVLQAGNTSCLLHYMPQAP-----	SAEPPLE	326	
QY	487	-----KSVLLELQOHL	497		
DB	327	WFPDLPSEPLCRGPILPQANL	349		

Search completed: March 5, 2003, 06:23:15
Job time : 98 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2003, 06:17:28 ; Search time 40 Seconds
(without alignments)
1735.589 Million cell updates/sec

Title: US-09-884-566A-2

Perfect score: 2650

Sequence: 1 MEVLESGEQSLQWDRKLSE.....LEGNETLKVELERRVNAVTF 521

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

1 number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2650	100.0	521	AAW37934	Mouse CAMP regulat
2	2062	77.8	674	ABG10287	Novel human diagn
3	1518	57.3	993	ABG08669	Novel human diagn
4	937.5	35.4	519	ABG66662	Human fat responsi
5	926	34.9	520	ABG66661	Rat fat responsi
6	887.5	33.5	247	ABG50249	Human transcriptio
7	786	29.7	363	AAAB42970	Human ORFX ORF2734
8	735.5	27.8	479	ABG10527	Novel human diagn
9	520	19.6	144	AAAB41941	Human ORFX ORF1705
10	479	18.1	244	AAAB54258	Human pancreatic c

11	435	16.4	518	22	ABG63292	Drosophila melanog
12	432	16.3	219	22	AAW371382	Human gene 15-enco
13	395	14.9	395	20	AAW39682	Secreted protein 1
14	395	14.9	395	20	AAW01635	Human PS214 derive
15	395	14.9	395	21	AAW00019	Human secreted pro
16	395	14.9	395	21	AAW64650	Human human homolo
17	394	14.9	395	22	AAW93311	Human protein HP10
18	392	14.8	395	22	AAW67820	Human leucine zipp
19	384.5	14.5	461	23	AAW99664	Human Kail (K1) pr
20	380	14.3	226	23	ABW72303	Rat protein isolat
21	358	13.5	252	22	AAU16574	Human novel secret
22	348	13.1	236	22	AAU16573	Human novel secret
23	348	13.1	239	22	AAU16153	Human novel secret
24	330.5	12.5	371	19	AAW75181	Fragment of human
25	326.5	12.3	371	21	AAW42805	Human ORFX ORF2569
26	326.5	12.3	371	21	AAW73331	HYPM clone 673766
27	273	10.3	123	23	ABP42366	Human ovarian anti
28	249	9.4	299	21	AAW43050	Human ORFX ORF2814
29	249	9.4	299	21	AAW25719	Human protein sequ
30	224	8.5	700	21	AAW83636	Endoplasmic reticu
31	211	8.0	670	21	AAW83635	Endoplasmic reticu
32	181.5	6.8	105	22	AAW91486	Human immune/haema
33	178.5	6.7	450	22	ABG19519	Novel human diagn
34	170.5	6.4	391	22	ABG19691	Novel human diagn
35	166	6.3	547	23	ABP41519	Human ovarian anti
36	162.5	6.1	423	22	ABW82497	Human protein sequ
37	160.5	6.1	172	22	ABW03512	Human musculoskele
38	160	6.0	44	22	AAW64792	Human brain expres
39	160	6.0	44	22	AAW77548	Human bone marrow
40	160	6.0	44	22	AAW37732	Peptide #11769 enc
41	160	6.0	44	23	ABG46575	Human peptide enco
42	158.5	6.0	271	22	AAE00656	Human activating t
43	152	5.7	334	15	AAW53644	c-jun gene gene pr
44	152	5.7	334	23	ABW57355	Mouse ischaemic co
45	150	5.7	589	21	AAW26762	Human Nrf2 protein

ALIGNMENTS

RESULT 1

AAW37934
ID AAW37934 standard; Protein; 521 AA.

XX AC AAW37934;

XX DT 10-AUG-1998 (first entry)

XX DE Mouse CAMP regulatory element binding protein.

XX KW Mouse CAMP regulatory element binding; mCREBa; antibody; inhibition;
mCREBa modulator.

XX OS Mus sp.

XX PN WO9812910-A2.

XX PD 02-APR-1998.

XX PF 26-SEP-1997; 97WO-US17288.

XX PR 27-SEP-1996; 96US-0721684.

XX PA (ICOS-) ICOS CORP.

XX PI Keegan KS;

XX DR WPI; 1998-230326/20.

XX N-PSDB; AAV29204.

XX PT Murine CAMP regulatory element binding protein A, mCREBa - useful
for systematic analysis of structure and function of mCREBa and
identification of those molecules with which it will react

XX Claim 1; Pages 30-31; 41pp; English.

PS This is the amino acid sequence of the mouse cAMP regulatory element

CC binding (mCREBA) protein. The products of the mCREBA such as

CC antibodies can be used for the recombinant production of the protein,

CC to identify novel genes encoding binding partner polypeptides for

CC mCREBA, for the systematic analysis of the structure and function of

CC mCREBA and identification of those molecules with which it will react,

CC to identify inhibitors of mCREBA binding to other natural binding

CC partners and to generate rodents that fail to express a functional

CC mCREBA or express a variant mCREBA, useful as models for studying the

CC activities of mCREBA and mCREBA modulators in vivo.

XX Sequence 521 AA;

SQ Query Match 100.0%; Score 2650; DB 19; Length 521;

Best Local Similarity 100.0%; Pred. No. 9.8e-190;

Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MEVLESGEVSQVLMQDRKLSFSGTEALMYHTHSELDDEFSQNVLGQLSDPFLSEK 60

1 MEVLESGEVSQVLMQDRKLSFSGTEALMYHTHSELDDEFSQNVLGQLSDPFLSEK 60

61 SESMEVEPPTSPAPLQAEHSYSLSEEPRTQSPFTHAATSDSFNDEVESEKWLSTEF 120

61 SESMEVEPPTSPAPLQAEHSYSLSEEPRTQSPFTHAATSDSFNDEVESEKWLSTEF 120

121 PSATIKKEPITEQPPGLVPSVLTITATSTPPEKESPLDMNAGDSSCQTLIPKIKLE 180

121 PSATIKKEPITEQPPGLVPSVLTITATSTPPEKESPLDMNAGDSSCQTLIPKIKLE 180

181 PHEVDQFLNFSPEASVDQLHLPTPPSHSDSGSLSPNPLHPSLSQAHSVRAMP 240

181 PHEVDQFLNFSPEASVDQLHLPTPPSHSDSGSLSPNPLHPSLSQAHSVRAMP 240

241 RGPALSTSPLLTAPHKLGSGPLVLTETEEKRTLVAEGYPIPTKLPTKSEKALKKIRR 300

241 RGPALSTSPLLTAPHKLGSGPLVLTETEEKRTLVAEGYPIPTKLPTKSEKALKKIRR 300

301 KIKNKISAQSRKKKKEYMDSLEKKVESCSTENLELRKKVEVLNTNRTLLQQLQTL 360

301 KIKNKISAQSRKKKKEYMDSLEKKVESCSTENLELRKKVEVLNTNRTLLQQLQTL 360

361 VMGKVRTCKLAGTGTCTCLMVVLCFAVAFGFFQGYGYPGATKVALPSQHPLEPYT 420

361 VMGKVRTCKLAGTGTCTCLMVVLCFAVAFGFFQGYGYPGATKVALPSQHPLEPYT 420

421 ASVVRGNLLIYEHAFLPSESSSPASTGELGWDGSSLLRASSGLEALPEVDLPFLIS 480

421 ASVVRGNLLIYEHAFLPSESSSPASTGELGWDGSSLLRASSGLEALPEVDLPFLIS 480

481 NETSLEKSVLLELQQLHVSSKLEGNETLKVVLELERRVNAIF 521

481 NETSLEKSVLLELQQLHVSSKLEGNETLKVVLELERRVNAIF 521

RESULT 2

ABG10287

ID ABG10287 standard; Protein; 674 AA.

XX AC ABG10287;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #10278.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO2001/75047-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS74474.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity -

XX Claim 20; SEQ ID No 40646; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. ABG00010-ABG30377 represent novel human

XX diagnostic amino acid sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 674 AA;

Query Match 77.8%; Score 2062; DB 22; Length 674;

Best Local Similarity 63.5%; Pred. No. 1.3e-145;

Matches 428; Conservative 16; Mismatches 34; Indels 196; Gaps 4;

QY 28 EALMYHTHSELDDEFSQNVLGQLSDPFLSEKSEMEVEPPTSPAPLQAEHSYSLSE 87

2 EALMYHTHSELDDEFSQNVLGQLSDPFLSEKSEMEVEPPTSPAPLQAEHSYSLSE 61

QY 88 EPRTPQPTTHAATSDSFNDEVESEKWLSTETPPEKESPLDMNAGDSSCQTLIPKIKLE 147

62 EPRTPQPTTHAATSDSFNDEVESEKWLSTETPPEKESPLDMNAGDSSCQTLIPKIKLE 121

QY 148 AISTPPEKESPLDMNAGDSSCQTLIPKIKLEPHEVDQFLNFSPEAS 196

122 AISTPPEKESPLDMNAGDSSCQTLIPKIKLEPHEVDQFLNFSPEAS 181

QY 197 ----- 196

182 CTCGSSGNSAIRIILAYSLALRVPTETTKKPRDLFGTKGREKKEET'SREKLYVAERKA 241

QY 197 -VDQLHLPTPPSHSDSGSLSPNPLHPSLSQAHSVRAMPRGPSALSTSPILLTAP 255

242 PVDHLHLPTPPSHSDSGSLSPNPLHPSLSQAHSVRAMPRAAPSALSSPSLLTAP 301

QY 256 HKLQGGPLVLTETEEKRTLVAEGYPIPTKLPLTKSEKALKKIRRKIKNKISAQSRKK 315

302 HKLQGGPLVLTETEEKRTLVAEGYPIPTKLPLTKSEKALKKIRRKIKNKISAQSRKK 361

Qy 316 KEYMSLEKKVESCSTENLELRKKVELENTN----- 347
 Db 362 KEYMSLEKKVESCSTENLELRKKVELENTNASHGELVIGVGKGRWDPLTSPHAG 421
 Qy 348 ----- 347
 Db 422 DQCLGSVOGWAGECTPRQLDGSRTTLISEEDPGSPMGGEALGAIVSNWGREALHAS 481
 Qy 348 -----RTLLQLOKLOKLOTLVMGKVSRTCKLA 372
 Db 482 RVGAVIQAPIAAGSDICLSDIRESATSSQTNISVTRTLQLOKLOKLOTLVMGKVSRTCKLA 541
 Qy 373 GTQGTCTLMVVVLCFAVAFSGFGGYPSPATKMLPSQHLSPYPTASVRSNLLIY 432
 Db 542 GTQGTCTLMVVVLCFAVAFSGFGGYPSPATKMLPSQHLSPYPTASVRSNLLIY 601
 Qy 433 BEHAPLEBSSSPASTGELGWDGSGSLRASSGLEALPEVDLPHEPLISNETSLEKSVILLE 492
 602 BEHSPPESSSPGAGELGGWDGSGSLLEV-SGLESPDVLPHFIISNETSLEKSVILLE 660
 Qy 493 LQOHLVSKLEGNE 506
 Db 661 LQOHLVSAKLEGNE 674

RESULT 3

ABG08669

ID ABG08669 standard; Protein; 993 AA.

XX AC ABG08669;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #8660.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-0508631.

XX PF 31-MAR-2000; 2000US-0540217.

XX PF 23-AUG-2000; 2000US-0649167.

XX PA (HYSEQ-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX PI WPI, 2001-639362/73.

XX DR N-PSDB; AAS72856.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

XX PT responsible for genetic disorders or other traits and to assess

XX PT biodiversity -

XX PS Claim 20; SEQ ID No 39028; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and

XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX CC and gene mapping, and in recombinant production of (II). The

XX CC polynucleotides are also used in diagnostics as expressed sequence tags

XX CC for identifying expressed genes. (I) is useful in gene therapy techniques

XX CC to restore normal activity of (II) or to treat disease states involving

XX CC (II). (II) is useful for generating antibodies against it, detecting or

XX CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIFO
 CC at ftp.wifo.int/pub/published_pct_sequences.

XX SQ Sequence 993 AA;

Query Match 57.3%; Score 1518; DB 22; Length 993;

Best Local Similarity 71.1%; Pred. No. 1.1e-104;

Matches 317; Conservative 24; Mismatches 55; Indels 50; Gaps 6;

Qy 75 PLI-----QAHSYSLSEPRTPQSPFTHAATSDSFNDESEKWLSTEPSSATIKKEP 129

Db 8 PIITILSHEVEQAMSLCVHPKT-----IVSQDEVESEKWLSTDFPSTSIKTEP 56

Qy 130 ITEOPPGCLVPSVTLTITAIPTPEKEESPLDMNAGGSSCOTLPKLEPHEVDQFLN 189

Db 57 ITDEPPGCLVPSVTLTITAIPTPEKEEPPLEMTGVDSQCOTIIPKLEPHEVDQFLN 116

Qy 190 FSPKEASVDQLHLPPTPPSHSSDSEGLSPNRLHPFLSQAHSPVRAMPSPSALSTS 249

Db 117 FSPKEAPVDHLHLPPTPPSHSGSDSEGLSPNRLHPFLSQAHSPVRAMPSPSALSSS 176

Qy 250 PLLTAPHKLGSGPLVLTBEEKRTLVAEGYPIPTKPLTKSEKALKKIRKIKKISAQ 309

Db 177 PLLTAPHKLGSGPLVLTBEEKRTLVAEGYPIPTKPLTKSEKALKKIRKIKKISAQ 236

Qy 310 ESRRKKKYMDSLEKKVESCSTENLELRKKVELENTNRTLLQLOKLOKLOTLVMGKVSRTC 369

Db 237 ESRRKKKYMDSLEKKVESCSTENLELRKKVELENTNRTLLQLOKLOKLOTLVMGKVSRTC 296

Qy 370 KLAGTQGTCTLMVVVLCFAVAFSGFGGYPSPATKMLPSQHLSPYPTASVRSRNL 429

Db 297 KLAGTQGTCTLMVVVLCFAVAFSGFGGYPSPATKMLPSQHLSPYPTASVEN---- 352

Qy 430 LIYEHAPLEBSSSPASTGELGWDGSGSLRASSGLE-----ALP----- 470

Db 353 -LEEMDKFLDKSTLPRLNQE-----EDESLENRQVTGSIEATINSLETKSGSPGPDGFTDK 406

Qy 471 -----EVDLPHEPLISNETSLEKSVLL 491

Db 407 FCQRYKEELVPFLKLFQSTKEGIL 432

RESULT 4

ABG66662

ID ABG66662 standard; Protein; 519 AA.

XX AC ABG66662;

XX DT 29-AUG-2002 (first entry)

XX DE Human fat responsive transcription factor FTF1.

XX KW Fatty acid regulated gene; polyunsaturated fatty acid disorder;

XX KW PUPA disorder; eczema; cardiovascular disorder; hypertriglyceridaemia;

XX KW dyslipidaemia; atherosclerosis; coronary artery disease;

XX KW cerebrovascular disease; peripheral vascular disease; inflammation;

XX KW sinusitis; asthma; pancreatitis; osteoarthritis; rheumatoid arthritis;

XX KW acne; body weight disorder; obesity; cachexia; anorexia;

XX KW psychiatric disorder; cancer; cystic fibrosis; pre-menstrual syndrome;

XX KW diabetes; diabetic complication; genetic polymorphism.

XX OS Homo sapiens.

XX

PN WO200240666-A2.
XX
PD 23-MAY-2002.
XX
PF 19-NOV-2001; 2001WO-CA01632.
XX
PR 17-NOV-2000; 2000US-248589P.
XX
PA (XENO-) XENON GENETICS INC.
XX
PI Winther MD, Goldberg YP, Knickle LC, Haardt M, Allen SJ, Ponton A;
PI De Antueno RJ, Jenkins DK, Nwaka SO;
XX
DR WPI; 2002-508327/54.
XX
DR N-PSDB; ABK94814.
XX
PT Novel isolated polypeptide segment encoded by fat regulated genes,
PT useful for diagnosing the presence of or a predisposition for a
PT disorder involving fatty acid regulated genes in a subject -
XX
XX Claim 14; Fig 15; 225pp; English.
XX
CC The invention describes an isolated polypeptide segment (I) whose genes
CC are fat regulated. (I) or the polynucleotide encoding it (II) are useful
CC for diagnosing the presence of or a predisposition for a disorder
CC involving fatty acid regulated genes in a subject. A composition
CC containing (I) or (II) is useful for treating a disorder involving fatty
CC acid regulated genes, where the disorder is selected from a
CC polynaturated fatty acid (PUFA) disorder, eczema, cardiovascular
CC disorders (such as hypertriglyceridaemia, dyslipidaemia, atherosclerosis,
CC coronary artery disease, cerebrovascular disease or peripheral vascular
CC disease), inflammation (such as sinusitis, asthma, pancreatitis,
CC osteoarthritis, rheumatoid arthritis or acne), body weight disorders
CC (such as obesity, cachexia or anorexia), psychiatric disorders, cancer,
CC cystic fibrosis, pre-menstrual syndrome, diabetes, and diabetic
CC complications. (I) or (II) is useful as research agent and materials for
CC discovery of treatments and diagnostics for a disease, particularly human
CC disease. (II) is useful for constructing nucleotide probes and primers,
CC for detecting genetic polymorphism, for detecting changes in the level of
CC expression of (II), and as a diagnostic tool. This is the amino acid
CC sequence of a protein regulated by fatty acids.
XX
SQ Sequence 519 AA;
Query Match 35.4%; Score 937.5; DB 23; Length 519;
Best Local Similarity 50.2%; Pred. No. 1.2e-61;
Matches 220; Conservative 52; Mismatches 117; Indels 49; Gaps 14;
35 HFSLELDFTSONVLGQLSDPPFLSEKSEMEVPSPTSPAPLIQAHSYLSLSEPRTPQSP 94
42 HFTENMEDFSNDLFFSFFDDPVDEKSPLDLMELD--SPTPGIQAHSYLSLSDSAPQSP 99
95 FTHAATSDSFNDEVESEKWLSTFPSPSATIKKEPITEEQF--PGLVPSVTLTITAITST- 151
100 LVPIKMETTTQD--ABHGAWALGHKLCSINVKQEQ-SPELPVDPPLAAPSAMAAAAWATT 156
152 -----PFKEEESPLDMNAGDSSCQTLPKIKLEPHVEDQFLNFSPEKASVDQLHPPT 205
157 PLGLGLSPLSR--LPFHPQAFGE---MTQLFVIAEPLVNFQTKVTPE----DLVQMPTT 207
206 PPSHSDSDSGSLSPNPLHPFSLQAHSVPVAMPGPSALSTSPLLTAPHKLQG-SGPL 264
208 PPSHSDSDSGSOSPR-SLPP-----SSVPRPMASSTAISSITPLTAPHKLQGTSGPL 260
265 VLTEEEKRTLVARGYPTPTKPLTKSEKALKIRKIKNKISQAESRRKKKGYMDSLEK 324
261 LLTEEEKRTLIAEGYPTPTKPLTKAEKALKVRKIKNKISQAESRRKKKGYVECLEK 320
325 KVESCSTENLEKKVEVLENTHTLLQLOKLOQLTMGVKVSRTCKLAGTGTGCLMVVV 384
321 KVFETFSNNELWKKVETLENARTLLQLOKLOQLTMGVKVSRTCKLAGTGTGCLMVAA 380
385 LCPWAFGSPFGQGYGYPYSATKNALPSQHPLEP--YTASVWVRNLLIYEEHAPLEES 442

Db 381 LCFVLVLGSLVPCLEPFGSSGQTV--KEDPLAAGVYVTAQMPRSLLFYDDGA----- 432
QY 443 SPASTGELGGWDRGSSLL 460
Db 433 -----GLWEDGRSTL 442
RESULT 5
ABG66661
ID ABG66661 standard; Protein; 520 AA.
XX AC ABG66661;
XX 29-AUG-2002 (first entry)
XX Rat fat responsive transcription factor FTF1.
XX Fatty acid regulated gene; polyunsaturated fatty acid disorder;
KW PUPA disorder; eczema; cardiovascular disorder; hypertriglyceridaemia;
KW dyslipidaemia; atherosclerosis; coronary artery disease;
KW cerebrovascular disease; peripheral vascular disease; inflammation;
KW sinusitis; asthma; pancreatitis; osteoarthritis; rheumatoid arthritis;
KW acne; body weight disorder; obesity; cachexia; anorexia;
KW psychiatric disorder; cancer; cystic fibrosis; pre-menstrual syndrome;
KW diabetes; diabetic complication; genetic polymorphism.
XX Rattus sp.
OS WO200240666-A2.
PN 23-MAY-2002.
PD 19-NOV-2001; 2001WO-CA01632.
XX 17-NOV-2000; 2000US-248589P.
XX (XENO-) XENON GENETICS INC.
XX Winther MD, Goldberg YP, Knickle LC, Haardt M, Allen SJ, Ponton A;
PI De Antueno RJ, Jenkins DK, Nwaka SO;
XX
XX WPI; 2002-508327/54.
XX N-PSDB; ABK94813.
XX Novel isolated polypeptide segment encoded by fat regulated genes,
PT useful for diagnosing the presence of or a predisposition for a
PT disorder involving fatty acid regulated genes in a subject -
XX
XX Claim 14; Fig 14; 225pp; English.
XX The invention describes an isolated polypeptide segment (I) whose genes
CC are fat regulated. (I) or the polynucleotide encoding it (II) are useful
CC for diagnosing the presence of or a predisposition for a disorder
CC involving fatty acid regulated genes in a subject. A composition
CC containing (I) or (II) is useful for treating a disorder involving fatty
CC acid regulated genes, where the disorder is selected from a
CC polynaturated fatty acid (PUFA) disorder, eczema, cardiovascular
CC disorders (such as hypertriglyceridaemia, dyslipidaemia, atherosclerosis,
CC coronary artery disease, cerebrovascular disease or peripheral vascular
CC disease), inflammation (such as sinusitis, asthma, pancreatitis,
CC osteoarthritis, rheumatoid arthritis or acne), body weight disorders
CC (such as obesity, cachexia or anorexia), psychiatric disorders, cancer,
CC cystic fibrosis, pre-menstrual syndrome, diabetes, and diabetic
CC complications. (I) or (II) is useful as research agent and materials for
CC discovery of treatments and diagnostics for a disease, particularly human
CC disease. (II) is useful for constructing nucleotide probes and primers,
CC for detecting genetic polymorphism, for detecting changes in the level of
CC expression of (II), and as a diagnostic tool. This is the amino acid
CC sequence of a protein regulated by fatty acids.
XX
SQ Sequence 520 AA;

Query Match 34.9%; Score 926; DB 23; Length 520;
Best Local Similarity 47.6%; Pred. No. 8.8e-61;
Matches 219; Conservative 49; Mismatches 142; Indels 50; Gaps 11;

QY 35 HFSELLDFSONVQLQSDPFLSEKSEMEVEPSPPTSPAPLQAEHSYSLSEEPRTQSP 94
DB 42 HFVENMEDFNDLFSFFDDPVLDEKSPLLDMELD--SPAPGQAEHSYSLSGDSAPQSP 99
QY 95 FTHAATSDSFNDEEVESEKWKYLSSTFPSPATIKKEPITEOPGLVPSVTLTITASTP-- 152
DB 100 LVFVVKMEDTTQD--MEHGAWALGNKLCSTIMVKQEOSPELVDPLAASSAMAATWATPPL 157
QY 153 ---PEKEESPIDMNAAGDSSCOTLPIKLEPHEVDQFLNFPKXASVDQLHLPLPTPPSS 209
DB 158 LGLSPISRLPIPHQAPGE---WTQLPVKAEPPMSQFL----KVTQEDLVQMPPTPPSS 210
QY 210 HSDSEGLSPNPLRHPLPFLSQAHSVPRAMPGRGPSALSTSPLLTAPHKLQG--SGPLVLTE 268
DB 211 HGSDDSGSOPR--SLPP-----SSPVRPMARSSTAISTSPLLTAPHKLQGTSGPLLLTE 263
269 EKRTLVAEGYPIPTKPLTKSEKALKIRKIKNKISAQSRKKKBYMDSLEKKVES 328
DB 264 EKRTLVAEGYPIPTKPLTKAEKALKVRREKIKNKISAQSRKKKBYVECLEKKVET 323
QY 329 CSTENLELRKKVELENTNRTLLQOLQKLOTLVMGKVSRTCKLACTQTCTCLMVVLGFA 388
DB 324 YISENNELWKKVETLETANRTLLQOLQKLOTLVTSKISRPYKMAATQTCTCLMVVAALCFV 383
QY 389 VAFGSFFOGYGPYPSPATKMAPLSQHPLSEPYTASVVRGNLLIYEHAPLEBSSSPASTG 448
DB 384 LVGLSLAPCLPAFSGSKTVKEDPVAADSVYAASQMPSSLLFYDDGA----- 431
QY 449 ELGHWG---RSSL-----LRASGLEALPEVDLPH 476
DB 432 --GSWEDHGRGALLVEPPEGWELKPGGPAEPQDHLRH 469

RESULT 5
ABB50249
ID ABB50249 standard; Protein; 247 AA.
XX ABB50249;
XX
XX
DT 05-FEB-2002 (first entry)
XX
DE
DE
XX
XX
Human transcription factor TRFX-100.
XX
XX
Human; transcription factor; TRFX; cell proliferative disease;
XX
XX
autoimmune disease; inflammation; neurological disease;
XX
XX
developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;
KW
KW
neuroprotective; antiinflammatory; gene therapy.
XX
OS
OS
Homo sapiens.
XX
XX
WO200172777-A2.
XX
XX
04-OCT-2001.
XX
XX
13-MAR-2001; 2001WO-US08117.
XX
XX
13-MAR-2000; 2000US-0188986.
XX
XX
(INCY-) INCYTE GENOMICS INC.
XX
XX
Hillman JL, Baughn MR, Yue H, Lal P, Lu DAM, Patterson C;
PI
PI
Azimzai Y, Bandman O, Tang YT, Mathur P, Shah P, Au-Young J;
PI
PI
Reddy R;
XX
XX
WPI; 2001-570896/64.
DR
DR
N-PSDB; ABA83073.
XX
XX
Novel transcription factor polypeptides, used to treat diseases
PT
PT
associated with altered activity and expression of TRFX, and to screen

for agents capable of modulating its activity -
Claim 1; Page 244; 327pp; English.

The present sequence is the protein sequence for a human transcription factor. The transcription factor and its coding sequence are useful in the diagnosis, treatment and prevention of diseases associated with altered expression of the transcription factor e.g. cell proliferative, autoimmune/inflammatory, neurological and developmental disorders. A number of specific disorders/diseases are given in the specification, including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS, allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic dermatitis, diabetes mellitus, emphysema, osteoarthritis, pancreatitis, Grave's disease, multiple sclerosis, osteoarthritis, psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's disease, stroke, and viral, bacterial, fungal and protozoal infections.

Sequence 247 AA;

Query Match 33.5%; Score 887.5; DB 22; Length 247;
Best Local Similarity 88.7%; Pred. No. 2.4e-58;
Matches 172; Conservative 8; Mismatches 13; Indels 1; Gaps 1;

QY 1 MEVLESGEOSVLQWDRKLSLSEPCGTEALMYHTHFSELLDFSONVQLQSDPFLSEK 60
DB 1 MEVLESGQGVQLQWDRKLSLSEPCGTEALMYHTHFSELLDFSONVQLQSDPFLSEK 60
QY 61 SESMEVEPSPSPAPLQAEHSYSLSEEPRTQSPPTTHAATSDSFNDEVESEKWLSTEF 120
DB 61 SVSMEVEPSPSPAPLQAEHSYSLSEEPRTQSPPTTHAATSDSFNDEVESEKWLSTDF 119
QY 121 PSATIKKEPITEOPGLVPSVTLTITASTPFEKEESPIDMNAAGDSSCOTLPIKLE 180
DB 120 PSTSIKTEPITDEPPGLVPSVTLTITASTPFEKEESPIDMNAAGDSSCOTLPIKLE 179
QY 181 PHEVDQFLNFPKE 194
DB 180 PHEVDQFLNFPKE 193

RESULT 7
AAB42970
ID AAB42970 standard; Protein; 363 AA.
XX AAB42970;
XX
XX
08-FEB-2001 (first entry)
XX
DE
DE
Human ORFX ORF2734 polypeptide sequence SEQ ID NO:5468.
XX
XX
Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW
KW
vulnary; antipsoriatic; antiparkinsonian; neuroprotective;
KW
KW
anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
KW
KW
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW
KW
hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW
KW
antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW
KW
antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW
KW
neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW
KW
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW
KW
cholesterol ester storage; systemic lupus erythematosus; infection;
KW
KW
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW
KW
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW
KW
bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW
KW
thrombosis; contraceptive.
XX
OS
OS
Homo sapiens.
XX
XX
WO200058473-A2.
XX
XX
05-OCT-2000.
XX
XX
31-MAR-2000; 2000WO-US08621.

```
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
XX 30-MAR-2000; 2000US-0540763.
PR
XX (CURA-) CURAGEN CORP.
XX
XX Shinketsu RA, Leach M;
XX
XX WPI; 2000-602362/57.
DR N-PSDB; AAC77179.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11; Page 4646-4647; 5507pp; English.
XX
XX AAC7446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX immunostimulant; cardiac; thrombolytic; coagulant; vasotrophic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX antithyroid; and antianemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 363 AA;
SQ
Query Match 29.7%; Score 786; DB 21; Length 363;
Best Local Similarity 52.8%; Pred. No. 1.6e-50;
Matches 187; Conservative 34; Mismatches 91; Indels 42; Gaps 13;
35 HFSELDDEFQNVGLQDLPFLSEKSESEVPSPTSPAPLIQAHSYSLSEPRTPSP 94
42 HFTENMEDFSNDLFSFDFDVLDEKSPLDMDLD--SPTPGIQAHSYSLSGSAPQSP 99
95 FTHAATSDSFNDESEKSYLSTPEPSPATIKKEPITEOP--PGLVPSVTLTITAIT- 151
100 LVPIKMDTQD--AEHGAVALGHKLCSINVQEQ-SPELPVDPPLAAPSAMAAAATT 156
152 -----PFKEEBSPLDMNAGDSSCQTLPKIKLEPHEVDQFLNFSKESAVDQLHLPPT 205
157 PLLGLSLPLSR--LPIPHQAPGE---MTQLPVIKAEPLEVNQFLKVTPE---DLVQMPPT 207
206 PPSHSDSBSGLSPNPLPFLPSLSQAHSVPVRAMPGPSALSTSPLTATPHKLOG-SGPL 264
208 PPSHSDSDSGSOSPR-SLPP-----SSPVRPMARSTALSSPLTATPHKLOGSGPL 260
265 VLTEEEKRTILVAEGYPTKPLPTKSEKALKKIRRIKKNISQAESRRKKKKEYMDSLEK 324
261 VLTEEEKRTILVAEGYPTKPLPTKSEKALKKIRRIKKNISQAESRRKKKKEYMDSLEK 320
325 KVESCSTENLELRKKVELENTNRTLQQLQKQTLVMGKVSRTCKLAGTGTGT 378
321 KVESCSTENLELRKKVETLENAN-SFSSGIQPL-----LCSLTIGLENPT 363
RESULT 8
```

ABG10527

ID ABG10527 standard; Protein; 479 AA.

XX AC ABG10527;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #10518.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX N-PSDB; AAS74714.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity -

XX Claim 20; SEQ ID No 40886; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (III). (III) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. ABG00010-ABG30377 represent novel human

XX diagnostic amino acid sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 479 AA;

Query Match 27.8%; Score 735.5; DB 22; Length 479;

Best Local Similarity 43.7%; Pred. No. 1.4e-46;

Matches 194; Conservative 48; Mismatches 95; Indels 107; Gaps 18;

QY 63 SMEVPSPTSPAPLIQAHSYSLSEPRTPSPFTHAATSD-SFND----- 106

DB 6 AMATSPA-AGPQLVQAAH-----PPDSTSIYTPGLISDFDPSGILGAPSQLMPPSVL 59

QY 107 -----EEVESEKWLSTPEPSPATIKKEPITE-----EQPPGL--VPSVTLTITAIS 150

DB 60 LLPAIEQLHCSRHYLL--FYS---HGNPHTKACVAFVGVGGPFLRGSFSLSCAONS 114

QY 151 TPPE-----KEESPLDNAGDSSCOTLIPKIKLEPHEVDOLFSPKEASVDQ 199
 Db 115 KSHFPNCAGQATSPREAPGEM-----TQLPVTKAEPLVNFQFLKVTPE-----DL 161
 QY 200 LHLPTTPSSSHSDSGSLSPNRLHPFSLQAHSPVRAMPGRGPSALSTSPLLITAPHKIQ 259
 Db 162 VQMPTTPSSSHSDSGSDGSPR-SLPP-----SSPVRPMARSSSTAISTSPLLITAPHKIQ 214
 QY 260 G-SGLVLVTEEEKRTLVAEGYDIPKPLPTKSEKALKIRKIKNKISQAQSRKKKEY 318
 Db 215 GTSGLVLLVTEEEKRTLVAEGYDIPKPLPTKAEKALKVRRKIKNKISQAQSRKKKEY 274
 QY 319 MDSLEKVKESCTENLELRKKVEVLNRTLLQQLQLOLQVLVGVSTCKLAGTQTGT 378
 Db 275 VECLKKVETFTSENELWKK-----LQKLQVLVTVTKISRPYKMAATQTGT 320
 QY 379 CLMVVLCFAVAFGFFQGYGYPYATKMLPSQHPLSEP--YTASVVRNLLIYEHA 436
 Db 321 CLMVAALCFVLVGLSLVPLCPFFSGQTV--KEDPLAADGVVTASQMPFSLFFYDDGA 378
 QY 437 PLEESSSPASTGELGWDGSGSLL 460
 Db 379 -----GLWEDGRSTL 388
 RESULT 9
 AAB41941
 ID AAB41941 standard; Protein; 144 AA.
 XX AC AAB41941;
 XX DT 08-FEB-2001 (first entry)
 XX DE Human ORFX ORF1705 polypeptide sequence SEQ ID NO:3410.
 XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipariatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antihypertensive;
 KW antiviral; antibacterial; antifungal; antitumor; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX OS Homo sapiens.
 XX PN WO200058473-A2.
 XX PD 05-OCT-2000.
 XX PF 31-MAR-2000; 2000WO-US08621.
 XX PR 31-MAR-1999; 99US-0127607.
 XX PR 02-APR-1999; 99US-0127636.
 XX PR 05-APR-1999; 99US-0127728.
 XX PR 30-MAR-2000; 2000US-0540763.
 XX FA (CURA-) CURAGEN CORP.
 XX PI Shimkets RA, Leach M;
 XX DR WPI; 2000-602362/57.
 XX DR N-ESDB; AAC76150.
 XX PT Novel nucleic acids and peptides derived from open reading frame X,
 XX useful for treating e.g. cancers, proliferative disorders,
 XX neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 2587; 5507pp; English.
 PS AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipariatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antitumor; antithyroid;
 CC antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX Sequence 144 AA;
 SQ
 Query Match 19.6%; Score 520; DB 21; Length 144;
 Best Local Similarity 89.4%; Pred. No. 3.4e-31;
 Matches 101; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MEVLESGQSVLQWDRKLSLSEPGETEALMYHTFSELLEFQSVNLQGLSDPFLSEK 60
 Db 1 MEVLESGQSVLQWDRKLSLSEPGETEALMYHTFSELLEFQSVNLQGLSDPFLSEK 60
 QY 61 SESMEVEPSPTSPAPLQAEHSYLSLSEPTQSPETHAATSDSFNDEVESEK 113
 Db 61 SVSMEVEPSPTSPAPLQAEHSYLSLSEPTQSPETHAATSDSFNDEVESEK 113
 RESULT 10
 AAB54258
 ID AAB54258 standard; Protein; 244 AA.
 XX AC AAB54258;
 XX DT 09-MAR-2001 (first entry)
 XX DE Human pancreatic cancer antigen protein sequence SEQ ID NO:710.
 XX KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
 KW detection; diagnosis; identification; cytostatic; neuroprotective;
 KW nootropic; immunomodulatory; relaxant; contraceptive; gynaecological;
 KW antiinflammatory; cardiac; gene therapy; chromosome mapping;
 KW linkage analysis; tissue identification; tissue typing; forensic;
 KW neural; immune system; muscular; reproductive; gastrointestinal;
 KW pulmonary; cardiovascular; renal; proliferative.
 XX OS Homo sapiens.
 XX PN WO2000055320-A1.
 XX PD 21-SEP-2000.
 XX PF 08-MAR-2000; 2000WO-US05989.
 XX PR 12-MAR-1999; 99US-0124270.
 XX FA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Ruben SM;
 XX WPI; 2000-579444/54.

DR N-PSDB; AAC99023.
XX
PT New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
XX cancer, or for use in assays for diagnosing a pathological condition -
XX
Claim 11; Page 1150-1151; 1379pp; English.
XX
XX
CC AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAC54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC synaerological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 244 AA;

Query Match 18.1%; Score 479; DB 21; Length 244;
Best Local Similarity 57.4%; Pred. No. 8.3e-28;
Matches 105; Conservative 21; Mismatches 39; Indels 18; Gaps 3;
QY 280 PIPKPLIKSEKALKKTRKKNKISAOESRRKKKEYMDSLEKKVESCSTENLELRKK 339
DB 1 PIPKPLIKSEKALKKTRKKNKISAOESRRKKKEYMDSLEKKVESCSTENLELRKK 60
QY 340 VELENTNRTLLQLOKQLTVLMGVKVSRTCKLAGTGTGTCMLVVLVCFVAFGFFQGYG 399
DB 61 VETLENANRTLLQLOKQLTVLMGVKVSRTCKLAGTGTGTCMLVVLVCFVAFGFFQGYG 120
QY 400 PYPATNMALPSQHPLEP--YTASVVRSNLLIYEHAPLESSESSPASTGELGGWDRGS 457
DB 121 EFSGSGQTV--KEDPLAAGVYTASQMPRSRLFFYDDG-----GLWEDGR 164
458 SLL 460
DB 165 STL 167

RESULT 11
AAB63292
ID AAB63292 standard; Protein; 518 AA.

XX AAB63292;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 16668.

XX Drosophila melanogaster.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX

PF 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX MPI; 2001-655860/75.
DR N-PSDB; ABL07395.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosure; SEQ ID NO 16668; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
XX sequences (ABLI01840-ABLI16175) and the encoded proteins
XX (AAB57737-ABB72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 518 AA;

Query Match 16.4%; Score 435; DB 22; Length 518;
Best Local Similarity 34.8%; Pred. No. 4.7e-24;
Matches 138; Conservative 50; Mismatches 112; Indels 96; Gaps 14;
QY 48 LQQLSDPFLSEKSEMEVPSPTSPAPLIQAEHSYSL-SSEPRTOQSPFTHAATSD---- 102
DB 129 ISRLSNPALN-----TSVADLTRSGSLQSLQAHQPHHGGSSHVVALEHF 176
QY 103 -----SPNDEVESEKWL-----STFPSTATIKKEPITEEPQPLVPS----- 141
DB 177 QLPQHLVDNDGSSVSLRDGMSPDICSDIETDESAIKDEPMSDSCPASPTSQASS 236
QY 142 ---VTITITAITSTP--FEKEESPDMNAGDSSCOTLIPK-----TKLEP 181
DB 237 QHQLSLNLAHLOSEMLFEPKHCGLLLTASNSN-NSLIKSQORQOILQODNLLMAAMEI 295
QY 182 HEVDQFLNFSPEASVDQLHLPTPPSS-HSSDSEGLSPNPRLHPF-----SLSQAH- 233
DB 296 KSEKQSTSNSSDKSHAHGYGIPLTFPSSLPDSDSEGNLSPEHLFAPLSPNATVISVANP 355
QY 234 ----SPVR-----AMPRGPSALST-----SPLLTAPHKLGSGPLV 265
DB 356 AGGESSVRVSRRTAASITRSSGSGSASGSGSTSTVTTRQPIHTPLISSQPK-GSTGLL 414
QY 266 LTBEEKRTLVAEGYPIPTKPLTKSEKALKKIRKKNKISAOESRRKKKEYMDSLEKK 325
DB 415 LTBEEKRTLVAEGYPIPTKPLTKSEKALKKIRKKNKISAOESRRKKKEYMDSLEKK 474
QY 326 VESCSTENLELRKKVELENTNRTLLQLOKQLTV 361
DB 475 VEILVTENHDYKRLGLEETNANLLSOLHKLQALV 510

RESULT 12
AAG71382
ID AAG71382 standard; Protein; 219 AA.
XX
XX AAG71382;
XX
XX 30-JUL-2001 (first entry)
XX
XX

Human gene 15-encoded secreted protein fragment, SEQ ID NO:234.

Human; secreted protein; proliferative disorder; cancer;
fetal abnormality; developmental abnormality; haematopoietic disorder;
immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
inflammation; allergy; neurological disorder; Alzheimer's disease;
Parkinson's disease; cognitive disorder; schizophrenia; asthma;
skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
cardiovascular disorder; angiogenic disorder; kidney disorder;
gastrointestinal disorder; pregnancy-related disorder; tumour;
endocrine disorder; infection; wound healing; vulnery;
cell culture; chemotaxis; food additive;
binding partner identification.

Homo sapiens.

WO200132674-A1.

10-MAY-2001.

25-OCT-2000; 2000WO-US293360.

29-OCT-1999; 99US-0162211.

30-JUN-2000; 2000US-0215138.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Komatsoulis GA, Young PE, Moore PA;

WPI; 2001-291051/30.

New nucleic acid molecule encoding a human secreted protein, useful for
preventing, treating or ameliorating medical conditions such as
rheumatoid arthritis, Alzheimer's disease and microbial infections -

Disclosure; Page 39; 581pp; English.

AAH31349-AAH31428 represent cDNAs corresponding to 26 human secreted
protein genes, and AAG71243-AAG71319 represent the proteins they encode.
AAG71320-AAG71403 represent human secreted protein fragments. The genes
and their corresponding secreted proteins are useful for preventing,
treating or ameliorating medical conditions, e.g., by protein or gene
therapy. Pathological conditions can be diagnosed by determining the
amount of the new protein in a sample or by determining the presence of
mutations in the new genes. Specific uses are described for each of the
52 genes, based on the tissues in which they are most highly expressed,
and include developing products for the diagnosis or treatment of
proliferative disorders, cancer, tumours, foetal and developmental
abnormalities, haematopoietic disorders, diseases of the immune system,
AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
allergies, neurological disorders (e.g., Alzheimer's disease,
Parkinson's disease), cognitive disorders, schizophrenia, asthma,
skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
cardiovascular disorders, angiogenic disorders, kidney disorders,
gastrointestinal disorders, pregnancy-related disorders, endocrine
disorders, and infections. The proteins can also be used to aid wound
healing and epithelial cell proliferation, to prevent skin aging due to
sunburn, to maintain organs before transplantation, for supporting cell
culture of primary tissues, to regenerate tissues, to identify their
cognate ligands or binding partners, and in chemotaxis, and can be used
as a food additive or preservative to modify storage properties.
Antibodies specific for a protein of the invention can be used in
alleviating symptoms associated with the disorders mentioned above, and
in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
immunosorbent assay (ELISA). The present sequence represents a human
secreted protein fragment referred to in the disclosure of the invention.

Sequence 219 AA;

Query Match 16.3%; Score 432; DB 22; Length 219;
Best Local Similarity 54.9%; Pred. No. 2.3e-24;
Matches 96; Conservative 22; Mismatches 39; Indels 18; Gaps 3;

QY 288 TKSEBKALKIRKIKNKISAEKSRKKKEYMDLSLEKKVESCSTENLELRKKVEVLENTN 347
Db 1 SKAEKALKURVKRKIKNKISAEKSRKKKEYVECKEUKVETPTSENELWKKVETLENAN 60
QY 348 RTLLQOLQKLTLMVNGKVRICKLAGTGTCTCLMVVILCFVAFVGFQGYGYPSPATKM 407
Db 61 RTLLQOLQKLTLMVNGKVRICKLAGTGTCTCLMVVILCFVAFVGFQGYGYPSPATKM 407
QY 408 ALPSOHPLSEPP--YTASVVRNRNLLIYEEHAPLEBSSSPASTGELGGWDRGSSLL 460
Db 121 V--KEDPLAADGVYATASQMPRSRLIFYDDGA-----GLWEDGGRSTL 159
RESULT 13
AAI59682
ID AAY59682 standard; Protein; 395 AA.
XX
AC AAY59682;
XX
DT 18-JAN-2000 (first entry)
XX
DE Secreted protein 108-009-5-0-A2-FL.
XX
KW Secreted protein; fingerprint identification technique;
KW chromosome mapping; human; hereditary disease; diagnosis; cancer;
KW hyperlipidaemia; cardiovascular; neurodegenerative disorder; therapy;
KW autoimmune disease; rheumatic disease; embryonic disorder; myopathy;
KW renal injury; amino aciduria; hypoglycaemia; male rat infertility;
KW hypertension.
XX
OS Homo sapiens.
XX
PN WO9940189-A2.
XX
PD 12-AUG-1999.
XX
PF 09-FEB-1999; 99WO-IB00282.
XX
PR 09-FEB-1998; 98US-0074121.
PR 13-APR-1998; 98US-0081563.
PR 10-AUG-1998; 98US-0096116.
PR 04-SEP-1998; 98US-0099273.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR WPI; 1999-600966/51.
XX
XX N-PSDB; AAZ40810.
XX
XX Extended cDNAs useful for expressing secreted proteins and to obtain
specific antibodies -
XX
XX Claim 10; Page 202-203; 244pp; English.
XX
XX This sequence represents a human secreted protein of the invention.
XX The extended cDNAs (or genomic DNAs obtainable from them) may be used to
XX prepare PCR primers and probes. These are useful for forensic matching or
XX positive identification by DNA sequencing. They may also be used in
XX alternative fingerprint identification techniques. Antibodies against the
XX proteins encoded by the extended cDNAs are useful in identification of
XX tissue types or cell species, as well as identifying tissue specific
XX soluble proteins. The sequences can be used for chromosome mapping and
XX identification of genes associated with hereditary diseases or drug
XX response. Signal sequences from the cDNAs can be used in construction of
XX secretion vectors. Other sequences derived from the extended cDNAs can be
XX used to clone upstream genomic DNA sequences including promoters. This is
XX in turn useful for identifying proteins that interact with promoter
XX sequences. Some of the proteins may be useful in diagnosing and treating
XX several disorders including, but not limited to: cancer, hyperlipidaemia,
XX cardiovascular and neurodegenerative disorders, autoimmune diseases, and
XX rheumatic diseases, embryonic disorders, hypertension, renal injury,
XX amino acidurias, hypoglycaemia, male rat infertility and myopathies.

XX SQ Sequence 395 AA;

Query Match 14.9%; Score 395; DB 20; Length 395;
 Best Local Similarity 34.0%; Pred. No. 3.1e-21;
 Matches 115; Conservative 49; Mismatches 108; Indels 66; Gaps 12;

Qy 135 PPGLVPSVTLTITATSTPFKEESPLDNAGDSSC--QTLP-----KIKLEPHEV--DQ 186
 Db 33 PPPEVPTVRL-----QEQGLQGMKSGDRCGLQESPEDFLKLFDIDPNEVYCS 82

Qy 187 FLNFPKESVDQLHLPPTPPSSHSDS-----EGSLSPNRLHPFSL 230
 Db 83 ASPGSDGISDSC--PSPAPRATSSPMLYEVVEAGALERMOGETGPNVGLISIQLD 141

Qy 231 QAHSVVRAMPGRPSALSTSPLLTAPHKLGSG-----PLVLTBEERKTLVA 276
 Db 142 Q--WSPAFWVP--DSCMVSELPDPAHAHILPRAGTVAPVPCVTLLPCQTLPLTDEERKLLGQ 199

Qy 277 EGYPIPTKLPLTKSEKALKIRKIKNKSIAQESRRKKKYYMDSLEKKVESCSTENLEL 336
 Db 200 EGVSLPSHLPLTKAERVLLKVRKIRNKQSAQDSRRKKKEYIDGLSRVAACSAQNQEL 259

Qy 337 RKKVEVLENTNLTLLQOLKQTLVGMKVSRTCKLAGTGTGTCMLVVLCFVAVAFSGFFQ 396
 Db 260 QKKVQELERHNLISLVAQLRQLQTL-----IAQTSNKA-AQTSTCVLILLFSLALILPSPFS 314

Qy 397 GYGPVPSATKMALPSQHPLSEPTASVVRNLLIYEE 434
 Db 315 PQSRPEAG-----SEDYQPHGVTSRNILTHKD 342

RESULT 14
 AAY01635
 ID AAY01635 standard; Protein; 395 AA.
 XX AC AAY01635;
 XX DT 22-JUN-1999 (first entry)
 XX DE Human PS214 derived polypeptide.
 XX KW PS214; prostate cancer; breast cancer; ovarian cancer;
 KW benign prostatic hyperplasia; BPH; prostatitis; ovarian cyst;
 KW prostatic intraepithelial neoplasia; PIN; atypical hyperplasia;
 KW fibroadenoma; cystic breast disease; ovarian cyst adenoma;
 KW ovarian endometrioma.
 XX OS Homo sapiens.
 XX WO9914357-A2.
 XX PD 25-MAR-1999.
 XX PF 18-SEP-1998; 98WO-US19496.
 XX PR 23-DEC-1997; 97US-0997074.
 XX PR 19-SEP-1997; 97US-0938383.
 XX PA (ABBO) ABBOTT LAB.
 XX PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;
 PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;
 PI Merchant BT, Roberts-Rapp L, Russell JC, Stroupe SD;
 XX WIPI; 1999-229547/19.
 XX New isolated PS214 polynucleotides and polypeptides used for, e.g.
 PT preventing or treating breast cancer
 XX Claim 23; Page 113-114; 116pp; English.
 XX Amino acid sequences AAY01635-38 represent PS214 derived polypeptides.

CC They are used for detecting, staging, preventing or treating diseases
 CC or conditions of the prostate, breast and ovary, particularly cancer.
 CC The polynucleotides are useful for detecting, diagnosing, staging,
 CC monitoring, prognosticating, in vivo imaging, preventing or treating,
 CC or determining predisposition to diseases or conditions of the prostate,
 CC breast and ovary such as prostate cancer, breast cancer and ovarian
 CC cancer. They can be used for conditions such as e.g. benign prostatic
 CC hyperplasia (BPH), prostatitis, prostatic intraepithelial neoplasia
 CC (PIN), atypical hyperplasia, fibroadenoma, cystic breast disease,
 CC ovarian cysts, ovarian cyst adenoma, or ovarian endometriomas.
 XX

SQ Sequence 395 AA;

Query Match 14.9%; Score 395; DB 20; Length 395;
 Best Local Similarity 34.0%; Pred. No. 3.1e-21;
 Matches 115; Conservative 49; Mismatches 108; Indels 66; Gaps 12;

Qy 135 PPGLVPSVTLTITATSTPFKEESPLDNAGDSSC--QTLP-----KIKLEPHEV--DQ 186
 Db 33 PPPEVPTVRL-----QEQGLQGMKSGDRCGLQESPEDFLKLFDIDPNEVYCS 82

Qy 187 FLNFPKESVDQLHLPPTPPSSHSDS-----EGSLSPNRLHPFSL 230
 Db 83 ASPGSDGISDSC--PSPAPRATSSPMLYEVVEAGALERMOGETGPNVGLISIQLD 141

Qy 231 QAHSVVRAMPGRPSALSTSPLLTAPHKLGSG-----PLVLTBEERKTLVA 276
 Db 142 Q--WSPAFWVP--DSCMVSELPDPAHAHILPRAGTVAPVPCVTLLPCQTLPLTDEERKLLGQ 199

Qy 277 EGYPIPTKLPLTKSEKALKIRKIKNKSIAQESRRKKKYYMDSLEKKVESCSTENLEL 336
 Db 200 EGVSLPSHLPLTKAERVLLKVRKIRNKQSAQDSRRKKKEYIDGLSRVAACSAQNQEL 259

Qy 337 RKKVEVLENTNLTLLQOLKQTLVGMKVSRTCKLAGTGTGTCMLVVLCFVAVAFSGFFQ 396
 Db 260 QKKVQELERHNLISLVAQLRQLQTL-----IAQTSNKA-AQTSTCVLILLFSLALILPSPFS 314

Qy 397 GYGPVPSATKMALPSQHPLSEPTASVVRNLLIYEE 434
 Db 315 PQSRPEAG-----SEDYQPHGVTSRNILTHKD 342

RESULT 15

AAG00019

ID AAG00019 standard; Protein; 395 AA.

XX AC AAG00019;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein #7.

XX KW Human; secreted protein; 5' EST; expressed sequence tag; cDNA isolation;
 gene therapy; chromosome mapping.

XX OS Homo sapiens.

XX PN EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX PA (GEST) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WIPI; 2000-500381/45.

XX DR N-PSDB; AAC00016.

XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for

PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Example 19; SEQ ID 14; 71pp + CD-ROM; English.
XX
CC The present sequence is a human secreted protein which shows extensive
CC homology to the bZIP family of transcription factors, and especially to
CC the human human protein. The full length cDNA encoding this protein was
CC obtained from a 5' EST using first and second strand synthesis
CC procedures. 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.

SQ	Sequence	395 AA;
Query Match	14.9%;	Score 395; DB 21; Length 395;
Best Local Similarity	34.0%;	Pred. No. 3.1e-21;
Matches 115; Conservative	49;	Mismatches 108; Indels 66; Gaps 12;
QY	135	PPGLVPSVTLTITAITPFKEESPDLNAGGDSGC--QTLIP-----KIKLEPHEV--DQ 186
Db	33	PPPEVPVTRL-----QEQLQGWKSGDRCCGLQSEPEDFLKFIDPNEVCSE 82
QY	187	FLNFSPKASVDQLHLPTPPSSHSDS-----EGSLSPNPRHLHPFSL 230
Db	83	ASPGSDSGISEDSCH-PDSPAPRATSPMLYEVVYEAALERMQGETGPNVGLISIQLD 141
QY	231	QAHSPVRAMPRGPSALSTSLTAPHKLGSG-----PLVLTTEEKRTIVA 276
Db	142	Q-WSPAFWVP--DSCWVSELPPDAHAHILFRAGTVAPVCTTLLPCQTLLTDEEKRLGQ 199
QY	277	EGYPIPTKLPLTKSEKALKIRKIKNKISAQESRRKKKEYMDSLEKKVSCSTENLEL 336
Db	200	EGVSLPSHLPLTKAEERVLKVRKIRNKQSAQDSRRKKKEYIDGLSRVACSAQNOEL 259
QY	337	RKKVEVLENTNRTLLOQLQKLVGMKVSRCTKLAGTCTCLMVVLCFAVAFGSFFQ 396
Db	260	QKVQELERHNISLVQLRQLQTL---IAQTSNKA-AQTSTCVLILFSLALILPSFS 314
QY	397	GYGYPSPATKWALPSQHPLEPYTASVVRSRNLLIYEE 434
Db	315	PFQSRPEAG-----SEDYQPHGVTSRNLTHKD 342

Search completed: March 5, 2003, 06:18:33
Job time : 43 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2003, 06:20:31 ; Search time 356 Seconds
(without alignments)
61.715 Million cell updates/sec

Title: US-09-884-566A-2

Perfect score: 2650

Sequence: 1 MEVLESEQSVLQWDRKLS.....LEGNETLKVELERRVATP 521

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

188354 seqs, 42170167 residues

1 number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/1/pubaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2650	100.0	521	9	US-09-884-566-2
2	479	18.1	244	10	US-09-925-297-710
3	380	14.3	226	9	US-09-866-050A-627
4	358	13.5	252	10	US-09-764-864-1527
5	348	13.1	236	10	US-09-764-864-1526
6	348	13.1	239	10	US-09-764-864-1106
7	330.5	12.5	371	9	US-09-981-876-259
8	330.5	12.5	371	9	US-09-148-545-259
9	160.5	6.1	172	10	US-09-764-877-1459
10	160	6.0	44	10	US-09-864-761-46079
11	152	5.7	334	9	US-09-165-522-14
12	150	5.7	589	9	US-09-962-855-4
13	146.5	5.5	331	9	US-10-051-989-10
14	146.5	5.5	331	9	US-09-861-097-10
15	145.5	5.5	192	10	US-09-772-656-2
16	145.5	5.5	192	10	US-09-772-656-6
17	145.5	5.5	192	10	US-09-772-656-10
18	144.5	5.5	581	9	US-09-962-855-2
19	143.5	5.4	2005	10	US-09-735-367B-3

20	143.5	5.4	2063	10	US-09-735-367B-2
21	142	5.4	230	9	US-10-050-185-27
22	140.5	5.3	1404	10	US-09-811-045A-1
23	139	5.2	1192	10	US-09-789-386-2
24	139	5.2	1192	10	US-09-758-140-6
25	139	5.2	1192	10	US-09-893-348-23
26	139	5.2	1192	10	US-09-972-599A-6
27	138	5.2	1781	10	US-09-738-877-3
28	138	5.2	2462	9	US-09-819-104A-5
29	134.5	5.1	941	12	US-10-124-557-14
30	134.5	5.1	1022	12	US-10-124-557-84
31	134.5	5.1	1038	12	US-10-124-557-74
32	134.5	5.1	1049	12	US-10-124-557-58
33	134.5	5.1	1140	12	US-10-124-557-104
34	134.5	5.1	1270	12	US-10-124-557-44
35	134.5	5.1	1311	12	US-10-124-557-42
36	134.5	5.1	1313	12	US-10-124-557-142
37	134.5	5.1	1314	12	US-10-124-557-50
38	134.5	5.1	1320	12	US-10-124-557-46
39	134.5	5.1	1320	12	US-10-124-557-60
40	134.5	5.1	1354	12	US-10-124-557-48
41	134.5	5.1	1361	12	US-10-124-557-40
42	134.5	5.1	1363	12	US-10-124-557-52
43	134.5	5.1	1404	12	US-10-124-557-2
44	134.5	5.1	1404	12	US-10-124-557-62
45	133	5.0	92	9	US-10-059-720-9

ALIGNMENTS

RESULT 1

US-09-884-566-2
; Sequence 2, Application US/09884566
; Publication No. US2002019712A1

GENERAL INFORMATION:

APPLICANT: Keegan, Kathleen S.
TITLE OF INVENTION: No. US2002019712A1el CREBA Isoform
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/884,566

FILING DATE: 19-Jun-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/407,715
FILING DATE: 28-Sep-1999
APPLICATION NUMBER: US/09/005,970
FILING DATE: 12-Jan-1998
APPLICATION NUMBER: US/08/721,684
FILING DATE: 27-Sep-1996

ATTORNEY/AGENT INFORMATION:

NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/33487

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 521 amino acids
TYPE: amino acid
TOPOLOGY: linear

```

; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-884-566-2

Query Match      100.0%; Score 2650; DB 9; Length 521;
Best Local Similarity 100.0%; Pred. No. 2.7e-168;
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEVLSEGEVSQVLDWRKLSLSEPGTEALMYHTHPSSELLDFSPQNVQLQLSDPFLSEK 60
DB 1 MEVLSEGEVSQVLDWRKLSLSEPGTEALMYHTHPSSELLDFSPQNVQLQLSDPFLSEK 60
QY 61 SSMEVEPSPTSPAPUIQAEHSVSLSEEPRTQSPFFHAATSDSFNDEEVESEKWLSTEF 120
DB 61 SSMEVEPSPTSPAPUIQAEHSVSLSEEPRTQSPFFHAATSDSFNDEEVESEKWLSTEF 120
QY 121 PSATIKKEPITEBPQPLVPSVTLTITATSTPPEKEESPMDNAGDSSCQTLIPKIKLE 180
DB 121 PSATIKKEPITEBPQPLVPSVTLTITATSTPPEKEESPMDNAGDSSCQTLIPKIKLE 180
QY 181 PHEVDQFLNFPKPEASVDQLHLPTPPSHSDSEGLSPNPRLHPFSLQAHSVPVRAMP 240
DB 181 PHEVDQFLNFPKPEASVDQLHLPTPPSHSDSEGLSPNPRLHPFSLQAHSVPVRAMP 240
QY 241 RGPSALSTSPLLTAPHKLOGSGPLVLTETBEKRTLVAEGYPIPTKLPTKSEKALKKIRR 300
DB 241 RGPSALSTSPLLTAPHKLOGSGPLVLTETBEKRTLVAEGYPIPTKLPTKSEKALKKIRR 300
QY 301 KIKNKISAESESRKKKYEYMDSEKVESCSSTENLELRKKEVLENTNRTLQLOKLQTL 360
DB 301 KIKNKISAESESRKKKYEYMDSEKVESCSSTENLELRKKEVLENTNRTLQLOKLQTL 360
QY 361 VMGKVRTCKLAGTGTGTCMLVVVLCFAVAFSGFFQGYGYPSPATKMLPSQHPLEPYT 420
DB 361 VMGKVRTCKLAGTGTGTCMLVVVLCFAVAFSGFFQGYGYPSPATKMLPSQHPLEPYT 420
QY 421 ASVVRNRLIYEHAFLPESSSPASTGELGWDGSSLLRASSGLEALPEVDLPHFLIS 480
DB 421 ASVVRNRLIYEHAFLPESSSPASTGELGWDGSSLLRASSGLEALPEVDLPHFLIS 480
QY 481 NETSLEKSVLLELQQLHVSLSKSGNETLKVLELERRVNTAF 521
DB 481 NETSLEKSVLLELQQLHVSLSKSGNETLKVLELERRVNTAF 521

RESULT 2
US-09-925-297-710
; Sequence 710, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 710
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (189)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (229)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-710
```

```

Query Match      18.1%; Score 479; DB 10; Length 244;
Best Local Similarity 57.4%; Pred. No. 1e-24;
Matches 105; Conservative 21; Mismatches 39; Indels 18; Gaps 3;

QY 280 PIPTKLPTKSEKALKKIRKIKNKISAESESRKKKYEYMDSEKVESCSSTENLELRKK 339
DB 1 PIPTKLPTKSEKALKKIRKIKNKISAESESRKKKYEYMDSEKVESCSSTENLELRKK 339
QY 340 VEVLNTNRTLQLOKLQTLVMGKVRTCKLAGTGTGTCMLVVVLCFAVAFSGFFQGYG 399
DB 61 VETLENANRTLQLOKLQTLVTNKRISRPYKMAATGTGTCMLVAALCFVLVLGSLVPCLP 120
QY 400 PYPSTATKMLPSQHPLESEPTASVVRNRLIYEHAFLPESSSPASTGELGWDGSS 457
DB 121 EFSGSGTQV--KEDPLADGVVTASQMPSSRLFFYDDGA-----GLWEDGR 164
QY 458 SLL 460
DB 165 STL 167

RESULT 3
US-09-866-050A-627
; Sequence 627, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-627

Query Match      14.3%; Score 380; DB 9; Length 226;
Best Local Similarity 47.6%; Pred. No. 3.4e-18;
Matches 90; Conservative 18; Mismatches 55; Indels 26; Gaps 3;

QY 300 RKIKNKISAESESRKKKYEYMDSEKVESCSSTENLELRKKEVLENTNRTLQLOKLQ 359
DB 1 RKIKNKISAESESRKKKYEYMDSEKVESCSSTENLELRKKEVLENTNRTLQLOKLQ 359
QY 360 LVMGKVRTCKLAGTGTGTCMLVVVLCFAVAFSGFFQGYGYPSPATKMLPSQHPLE 419
DB 61 LVTSKISRYPYKMAATGTGTCMLVAALCFVLVLGSLAPCLPAFSSGSKTVKEDVPA 120
QY 420 TASVVRNRLIYEHAFLPESSSPASTGELGWD---RGSSL-----LRASSGLE 467
DB 121 AASQMPSSRLFFYDDGA-----GSWEDGHRGALLPVEPPEGWELKPGGP 166
QY 468 ALPEVDLPH 476
DB 167 PRPDHLRH 175

RESULT 4
US-09-764-864-1527
; Sequence 1527, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```

```

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1527
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (188)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1527

Query Match      13.5%; Score 358; DB 10; Length 252;
Best Local Similarity 36.5%; Pred. No. 1.1e-16;
Matches 95; Conservative 42; Mismatches 81; Indels 42; Gaps 6

QY 221 NPLRHPSLSQAHSPVRAMPGPSALSTSPLLTAPHKLGSGPLVLTEEEKRTLVAEGYP 280
Db 10 NSRVDPRVRAHAHILPRAGTVAPVCTT-----LLPCQTLFTDEKRLGQEGYS 60

QY 281 IPTKLPTKSEKALKKTRRKIKNKISQAESRRKKKEYWDSLEKKVSCSTENLELRKKV 340
Db 61 LPSHLPLTKABEERVLKVKRRKIRNKQSAQDSRRRKKEYIDGLSEKRVACSAQNQELQKKV 120

QY 341 EVLENTNRTLQOLQKLOTLVGMKYSRCKLAGTGTGTCMLVVLCFAVAFSGFFQGYGP 400
Db 121 QELERHNSISLVAQRQLQTL----IAQTSNKA-AQTSTCVLILFSLALILPSPFPQS 175

QY 401 YPSATKMAPLPSQHPLESEPTASVRSRNLIIYEE-----HAPLEE-----SS 442
Db 176 RPEAG-----SEYQPHKVTSRNIIHTKDVNTLETQVVSERLEPPGAKDANGS 225

QY 443 SPASTGELGGWDGSSLLRA 462
Db 226 TRTLEKMGKGRPRSPGRIRS 245

RESULT 5
US-09-764-864-1526
; Sequence 1526, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1526
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1526

Query Match      13.1%; Score 348; DB 10; Length 236;
Best Local Similarity 45.0%; Pred. No. 4.8e-16;
Matches 77; Conservative 33; Mismatches 46; Indels 15; Gaps 3

QY 264 LVLTEEEKRTLVAEGYPPTKLPTKSEKALKKTRRKIKNKISQAESRRKKKEYWDSLE 323
Db 28 LFLTDEEKRLLGQEGVSLPFLPTKABEERVLKVKRRKIRNKQSAQDSRRRKKEYIDGLE 87

QY 324 KKVESCSTENLELRKKVEVLENTNRTLQOLQKLOTLVGMKYSRCKLAGTGTGTCMLVV 383
Db 88 SRVACSAQNQELQKKVQELERHNSISLVAQRQLQTL----IAQTSNKA-AQTSTCVLIL 142

```

1 PRIOR APPLICATION NUMBER: 60/047,615
2 PRIOR FILING DATE: 1997-05-23
3 PRIOR APPLICATION NUMBER: 60/047,600
4 PRIOR FILING DATE: 1997-05-23
5 PRIOR APPLICATION NUMBER: 60/047,597
6 PRIOR FILING DATE: 1997-05-23
7 PRIOR APPLICATION NUMBER: 60/047,502
8 PRIOR FILING DATE: 1997-05-23
9 PRIOR APPLICATION NUMBER: 60/047,633
10 PRIOR FILING DATE: 1997-05-23
11 PRIOR APPLICATION NUMBER: 60/047,583
12 PRIOR FILING DATE: 1997-05-23
13 PRIOR APPLICATION NUMBER: 60/047,617
14 PRIOR FILING DATE: 1997-05-23
15 PRIOR APPLICATION NUMBER: 60/047,618
16 PRIOR FILING DATE: 1997-05-23
17 PRIOR APPLICATION NUMBER: 60/047,503
18 PRIOR FILING DATE: 1997-05-23
19 PRIOR APPLICATION NUMBER: 60/047,592
20 PRIOR FILING DATE: 1997-05-23
21 PRIOR APPLICATION NUMBER: 60/047,581
22 PRIOR FILING DATE: 1997-05-23
23 PRIOR APPLICATION NUMBER: 60/047,584
24 PRIOR FILING DATE: 1997-05-23
25 PRIOR APPLICATION NUMBER: 60/047,500
26 PRIOR FILING DATE: 1997-05-23
27 PRIOR APPLICATION NUMBER: 60/047,587
28 PRIOR FILING DATE: 1997-05-23
29 PRIOR APPLICATION NUMBER: 60/047,492
30 PRIOR FILING DATE: 1997-05-23
31 PRIOR APPLICATION NUMBER: 60/047,598
32 PRIOR FILING DATE: 1997-05-23
33 PRIOR APPLICATION NUMBER: 60/047,613
34 PRIOR FILING DATE: 1997-05-23
35 PRIOR APPLICATION NUMBER: 60/047,582
36 PRIOR FILING DATE: 1997-05-23
37 PRIOR APPLICATION NUMBER: 60/047,596
38 PRIOR FILING DATE: 1997-05-23
39 PRIOR APPLICATION NUMBER: 60/047,612
40 PRIOR FILING DATE: 1997-05-23
41 PRIOR APPLICATION NUMBER: 60/047,632
42 PRIOR FILING DATE: 1997-05-23
43 PRIOR APPLICATION NUMBER: 60/047,601
44 PRIOR FILING DATE: 1997-05-23
45 PRIOR APPLICATION NUMBER: 60/043,580
46 PRIOR FILING DATE: 1997-04-11
47 PRIOR APPLICATION NUMBER: 60/043,568
48 PRIOR FILING DATE: 1997-04-11
49 PRIOR APPLICATION NUMBER: 60/043,314
50 PRIOR FILING DATE: 1997-04-11
51 PRIOR APPLICATION NUMBER: 60/043,569
52 PRIOR FILING DATE: 1997-04-11
53 PRIOR APPLICATION NUMBER: 60/043,311
54 PRIOR FILING DATE: 1997-04-11
55 PRIOR APPLICATION NUMBER: 60/043,671
56 PRIOR FILING DATE: 1997-04-11
57 PRIOR APPLICATION NUMBER: 60/043,674
58 PRIOR FILING DATE: 1997-04-11
59 PRIOR APPLICATION NUMBER: 60/043,669
60 PRIOR FILING DATE: 1997-04-11
61 PRIOR APPLICATION NUMBER: 60/043,312
62 PRIOR FILING DATE: 1997-04-11
63 PRIOR APPLICATION NUMBER: 60/043,313
64 PRIOR FILING DATE: 1997-04-11
65 PRIOR APPLICATION NUMBER: 60/043,672
66 PRIOR FILING DATE: 1997-04-11
67 PRIOR APPLICATION NUMBER: 60/043,315
68 PRIOR FILING DATE: 1997-04-11
69 PRIOR APPLICATION NUMBER: 60/048,974
70 PRIOR FILING DATE: 1997-06-06
71 PRIOR APPLICATION NUMBER: 60/056,886
72 PRIOR FILING DATE: 1997-08-22
73 PRIOR APPLICATION NUMBER: 60/056,877
74 PRIOR FILING DATE: 1997-08-22
75 PRIOR APPLICATION NUMBER: 60/056,889
76 PRIOR FILING DATE: 1997-08-22
77 PRIOR APPLICATION NUMBER: 60/056,893
78 PRIOR FILING DATE: 1997-08-22
79 PRIOR APPLICATION NUMBER: 60/056,630
80 PRIOR FILING DATE: 1997-08-22
81 PRIOR APPLICATION NUMBER: 60/056,878
82 PRIOR FILING DATE: 1997-08-22
83 PRIOR APPLICATION NUMBER: 60/056,662
84 PRIOR FILING DATE: 1997-08-22
85 PRIOR APPLICATION NUMBER: 60/056,872
86 PRIOR FILING DATE: 1997-08-22
87 PRIOR APPLICATION NUMBER: 60/056,882
88 PRIOR FILING DATE: 1997-08-22
89 PRIOR APPLICATION NUMBER: 60/056,637
90 PRIOR FILING DATE: 1997-08-22
91 PRIOR APPLICATION NUMBER: 60/056,903
92 PRIOR FILING DATE: 1997-08-22
93 PRIOR APPLICATION NUMBER: 60/056,888
94 PRIOR FILING DATE: 1997-08-22
95 PRIOR APPLICATION NUMBER: 60/056,879
96 PRIOR FILING DATE: 1997-08-22
97 PRIOR APPLICATION NUMBER: 60/056,880
98 PRIOR FILING DATE: 1997-08-22
99 PRIOR APPLICATION NUMBER: 60/056,894
100 PRIOR FILING DATE: 1997-08-22
101 PRIOR APPLICATION NUMBER: 60/056,911
102 PRIOR FILING DATE: 1997-08-22
103 PRIOR APPLICATION NUMBER: 60/056,636
104 PRIOR FILING DATE: 1997-08-22
105 PRIOR APPLICATION NUMBER: 60/056,874
106 PRIOR FILING DATE: 1997-08-22
107 PRIOR APPLICATION NUMBER: 60/056,910
108 PRIOR FILING DATE: 1997-08-22
109 PRIOR APPLICATION NUMBER: 60/056,864
110 PRIOR FILING DATE: 1997-08-22
111 PRIOR APPLICATION NUMBER: 60/056,631
112 PRIOR FILING DATE: 1997-08-22
113 PRIOR APPLICATION NUMBER: 60/056,845
114 PRIOR FILING DATE: 1997-08-22
115 PRIOR APPLICATION NUMBER: 60/056,892
116 PRIOR FILING DATE: 1997-08-22
117 PRIOR APPLICATION NUMBER: 60/047,595
118 PRIOR FILING DATE: 1997-05-23
119 PRIOR APPLICATION NUMBER: 60/057,761
120 PRIOR FILING DATE: 05-Sep-1997
121 PRIOR APPLICATION NUMBER: 60/047,599
122 PRIOR FILING DATE: 1997-05-23
123 PRIOR APPLICATION NUMBER: 60/047,588
124 PRIOR FILING DATE: 1997-05-23
125 PRIOR APPLICATION NUMBER: 60/047,585
126 PRIOR FILING DATE: 1997-05-23
127 PRIOR APPLICATION NUMBER: 60/047,586
128 PRIOR FILING DATE: 1997-05-23
129 PRIOR APPLICATION NUMBER: 60/047,590
130 PRIOR FILING DATE: 1997-05-23
131 PRIOR APPLICATION NUMBER: 60/047,594
132 PRIOR FILING DATE: 1997-05-23
133 PRIOR APPLICATION NUMBER: 60/047,589
134 PRIOR FILING DATE: 1997-05-23
135 PRIOR APPLICATION NUMBER: 60/047,593
136 PRIOR FILING DATE: 1997-05-23
137 PRIOR APPLICATION NUMBER: 60/047,614
138 PRIOR FILING DATE: 1997-05-23
139 PRIOR APPLICATION NUMBER: 60/043,578
140 PRIOR FILING DATE: 1997-04-11
141 PRIOR APPLICATION NUMBER: 60/043,576
142 PRIOR FILING DATE: 1997-04-11
143 PRIOR APPLICATION NUMBER: 60/047,501
144 PRIOR FILING DATE: 1997-05-23
145 PRIOR APPLICATION NUMBER: 60/043,670
146 PRIOR FILING DATE: 1997-04-11

```
; PRIOR APPLICATION NUMBER: 60/056,632
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,876
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,881
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,909
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,875
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,862
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,887
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,908
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/057,650
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/056,884
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 259
; LENGTH: 371

Query Match          12.5%  Score 330.5; DB 9; Length 371;
Best Local Similarity 30.7%; Pred. No. 1.3e-14;
Matches 116; Conservative 56; Mismatches 123; Indels 83; Gaps 15;

Qy 150 STPFKEKSPMDNAGDSCTQIPKIKLEPHEVDQFLNFPKESASVDQLHLP-PTPPS 208
Db 25 TAPDEAVRAPLDW-----ALP-----LSEVPSDWEVDLLCSLLSPPA 62

Qy 209 SHSDSEGLSPNPL--HPFSLQAHPVRAMPGRGALSTSLPTAPH-----KQSGS 262
Db 63 SLNLSL-----SSNCLVHHDTYSLPRETVSMDLSESCRKEGTQMTQHMBELAEQIEA 118

Qy 263 PLVLTTEERKTLVARGYPITPKLPTKSEKALKIRKKNKISAOESRRKKKEYMDSL 322
Db 119 RLVLDEKSLLEKGLILPETLPKTEEQELKKRKRIRKRAQESRRKKKYVVGSL 178

Qy 323 EKKVSCSTENLELRKKEVLENTNRTLLQQLKQLTVMGVKSVRTCKLAGTGTCTCLMV 382
Db 179 ESRVLKYTAQNWELQNLQLEEQNLSDLDQLRKLQAMVIEISNKT-----SSSSTCLIV 233

Qy 383 VVLCFAVAFGFFQGYGYGPSATKVALPSQH-----PLSEPYTASVVRSRNLLIY 432
Db 234 LLVSFCLLLVP-----AMYSSTRGSLPAEHGVLQRQLRALPSEDYPQLEL----- 279

Qy 433 BEHAPLESSPA-STGELGWDGSS--LLRASSGLEAL-----PEVDLPHELI 479
Db 280 ----PALQSEVPKDSHTQ---WLDGSDCVLQAGNNTSCLLHYMPQASBPPEWPFDDL 332

Qy 480 SNETSLEKSVLLEQOHL 497
Db 333 SSE-PLCRGPILPLQANL 349

RESULT 8
US-09-148-545-259
; Sequence 259, Application US/09148545
; Publication No. US20030027132A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2003, 06:17:30 ; Search time 43 Seconds
(without alignments)
356.496 Million cell updates/sec

Title: US-09-884-566A-2
Perfect score: 2650
Sequence: 1 MEVLESGSQSVLQWRKLSLSEPCGTEALMYHVFSELLDFEQNVLGQLLSDPFLSEK 521

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
al number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2650	100.0	521	2	US-08-721-684C-2
2	2650	100.0	521	2	US-09-005-970-2
3	2650	100.0	521	4	US-09-407-715-2
4	395	14.9	395	4	US-09-247-155-113
5	146.5	5.5	327	2	US-07-684-965-6
6	146.5	5.5	327	4	US-09-215-098-5
7	146.5	5.5	327	4	US-09-252-658-6
8	146.5	5.5	331	1	US-08-094-533B-10
9	146.5	5.5	331	1	US-08-276-860A-10
10	146.5	5.5	331	1	US-08-444-393-10
11	146.5	5.5	331	1	US-08-799-913-10
12	146.5	5.5	331	2	US-08-711-893-10
13	146.5	5.5	331	2	US-09-150-200-10
14	146.5	5.5	331	3	US-09-150-201-10
15	146.5	5.5	331	4	US-09-452-370-10
16	146.5	5.5	331	4	US-09-461-649-10
17	146.5	5.5	331	5	PCT-US95-15353-39
18	144.5	5.5	392	4	US-09-416-050A-2
19	144.5	5.5	392	4	US-09-664-800-2
20	144.5	5.5	392	4	US-09-665-309-2
21	144.5	5.5	392	4	US-09-661-569-2
22	143.5	5.4	295	2	US-08-343-443B-107
23	142	5.4	230	4	US-09-387-418A-27
24	140.5	5.3	1404	4	US-08-801-308-1
25	137.5	5.2	331	5	PCT-US94-08119-10
26	137.5	5.2	331	5	PCT-US94-08120-10
27	137.5	5.2	331	5	PCT-US94-12913A-10

28	134.5	5.1	941	4	US-07-757-022B-14	Sequence 14, Appl
29	134.5	5.1	1022	4	US-07-757-022B-84	Sequence 84, Appl
30	134.5	5.1	1038	4	US-07-757-022B-74	Sequence 74, Appl
31	134.5	5.1	1049	4	US-07-757-022B-58	Sequence 58, Appl
32	134.5	5.1	1140	4	US-07-757-022B-104	Sequence 104, Appl
33	134.5	5.1	1270	4	US-07-757-022B-44	Sequence 44, Appl
34	134.5	5.1	1311	4	US-07-757-022B-42	Sequence 42, Appl
35	134.5	5.1	1313	4	US-07-757-022B-142	Sequence 142, Appl
36	134.5	5.1	1314	4	US-07-757-022B-50	Sequence 50, Appl
37	134.5	5.1	1320	4	US-07-757-022B-46	Sequence 46, Appl
38	134.5	5.1	1320	4	US-07-757-022B-60	Sequence 60, Appl
39	134.5	5.1	1354	4	US-07-757-022B-48	Sequence 48, Appl
40	134.5	5.1	1361	4	US-07-757-022B-40	Sequence 40, Appl
41	134.5	5.1	1363	4	US-07-757-022B-52	Sequence 52, Appl
42	134.5	5.1	1404	4	US-07-757-022B-2	Sequence 2, Appl
43	134.5	5.1	1404	4	US-07-757-022B-62	Sequence 62, Appl
44	133	5.0	92	4	US-09-299-495F-9	Sequence 9, Appl
45	133	5.0	93	2	US-08-690-011A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-08-721-684C-2
; Sequence 2, Application US/08721684C
; Patent No. 5854016
; GENERAL INFORMATION:
; APPLICANT: Keegan, Kathleen S.
; TITLE OF INVENTION: No. 5854016el CREBa Isoform
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/721.684C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/33487
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-721-684C-2

Query Match 100.0%; Score 2650; DB 2; Length 521;
Best Local Similarity 100.0%; Pred. No. 5.8e-215;
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEVLESGSQSVLQWRKLSLSEPCGTEALMYHVFSELLDFEQNVLGQLLSDPFLSEK 60		
Db	1	MEVLESGSQSVLQWRKLSLSEPCGTEALMYHVFSELLDFEQNVLGQLLSDPFLSEK 60		
Qy	61	SESMEVEPSPTSPAPLIAQHSYLSSEPRQSPPTHAATSDSFNDEEVESEKWLSTEF 120		
Db	61	SESMEVEPSPTSPAPLIAQHSYLSSEPRQSPPTHAATSDSFNDEEVESEKWLSTEF 120		

QY 121 PSATIKKEPITEEQPPGLVPSVTLTITATISTPPEKEESPLDMNAGDSSCQTLIPKIKLE 180
DB 121 PSATIKKEPITEEQPPGLVPSVTLTITATISTPPEKEESPLDMNAGDSSCQTLIPKIKLE 180
QY 181 PHEVDOFLNFSPEASVDQLHLPPTPPSSSHSDSEGSLSPNRLHPFSLQAHSVPVRAMP 240
DB 181 PHEVDOFLNFSPEASVDQLHLPPTPPSSSHSDSEGSLSPNRLHPFSLQAHSVPVRAMP 240
QY 241 RGPSALSTSPLLTAPHKLGSGPLVLTBEEKRTLVAEGYPIPTKPLTKSEKALKKIRR 300
DB 241 RGPSALSTSPLLTAPHKLGSGPLVLTBEEKRTLVAEGYPIPTKPLTKSEKALKKIRR 300
QY 301 KIKNKISAESRRKKKEYMDSLEKKVSCSTENLELRKKVEVLENTNRTLLQOLKQL 360
DB 301 KIKNKISAESRRKKKEYMDSLEKKVSCSTENLELRKKVEVLENTNRTLLQOLKQL 360
QY 361 VMKVSRTCKLAGTGTGTCMLVVVLCFAVAFSGFFQGYGYPYPSATKMALPSQHPLEPYT 420
DB 361 VMKVSRTCKLAGTGTGTCMLVVVLCFAVAFSGFFQGYGYPYPSATKMALPSQHPLEPYT 420
QY 421 ASVVRSRNLLIYEEHAPLEBESSPASTGBLGWDGSSLLRASSGLEALPEVDLPHFLIS 480
DB 421 ASVVRSRNLLIYEEHAPLEBESSPASTGBLGWDGSSLLRASSGLEALPEVDLPHFLIS 480
QY 481 NETSLEKSVLLELQOHLVSSKLEGNETLKVLELRRVNTAF 521
DB 481 NETSLEKSVLLELQOHLVSSKLEGNETLKVLELRRVNTAF 521

RESULT 2

US-09-005-970-2
; Sequence 2, Application US/09005970
; Patent No. 5959079
; GENERAL INFORMATION:
; APPLICANT: Keegan, Kathleen S.
; TITLE OF INVENTION: No. 5959079el CREBa Isoform
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; City: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/721,684
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/33487
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-005-970-2
Query Match . 100.0%; Score 2650; DB 2; Length 521;

Best Local Similarity 100.0%; Pred. No. 5.8e-215;
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEVLSEGSQSVLOWDRKLSLSEPGETALMYHTFSELLDFEQNVQLQLSDPFLSEK 60
DB 1 MEVLSEGSQSVLOWDRKLSLSEPGETALMYHTFSELLDFEQNVQLQLSDPFLSEK 60
QY 61 SESMEVEPSPTSPALIOAHSYSLSSEPRTPSTPHTAATSDSPNDEEVESEKWLSTEF 120
DB 61 SESMEVEPSPTSPALIOAHSYSLSSEPRTPSTPHTAATSDSPNDEEVESEKWLSTEF 120
QY 121 PSATIKKEPITEEQPPGLVPSVTLTITATISTPPEKEESPLDMNAGDSSCQTLIPKIKLE 180
DB 121 PSATIKKEPITEEQPPGLVPSVTLTITATISTPPEKEESPLDMNAGDSSCQTLIPKIKLE 180
QY 181 PHEVDOFLNFSPEASVDQLHLPPTPPSSSHSDSEGSLSPNRLHPFSLQAHSVPVRAMP 240
DB 181 PHEVDOFLNFSPEASVDQLHLPPTPPSSSHSDSEGSLSPNRLHPFSLQAHSVPVRAMP 240
QY 241 RGPSALSTSPLLTAPHKLGSGPLVLTBEEKRTLVAEGYPIPTKPLTKSEKALKKIRR 300
DB 241 RGPSALSTSPLLTAPHKLGSGPLVLTBEEKRTLVAEGYPIPTKPLTKSEKALKKIRR 300
QY 301 KIKNKISAESRRKKKEYMDSLEKKVSCSTENLELRKKVEVLENTNRTLLQOLKQL 360
DB 301 KIKNKISAESRRKKKEYMDSLEKKVSCSTENLELRKKVEVLENTNRTLLQOLKQL 360
QY 361 VMKVSRTCKLAGTGTGTCMLVVVLCFAVAFSGFFQGYGYPYPSATKMALPSQHPLEPYT 420
DB 361 VMKVSRTCKLAGTGTGTCMLVVVLCFAVAFSGFFQGYGYPYPSATKMALPSQHPLEPYT 420
QY 421 ASVVRSRNLLIYEEHAPLEBESSPASTGBLGWDGSSLLRASSGLEALPEVDLPHFLIS 480
DB 421 ASVVRSRNLLIYEEHAPLEBESSPASTGBLGWDGSSLLRASSGLEALPEVDLPHFLIS 480
QY 481 NETSLEKSVLLELQOHLVSSKLEGNETLKVLELRRVNTAF 521
DB 481 NETSLEKSVLLELQOHLVSSKLEGNETLKVLELRRVNTAF 521

RESULT 3
US-09-407-715-2
; Sequence 2, Application US/09407715
; Patent No. 6248532
; GENERAL INFORMATION:
; APPLICANT: Keegan, Kathleen S.
; TITLE OF INVENTION: No. 6248532el CREBa Isoform
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; City: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/407,715
; FILING DATE: 28-Sep-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/005,970
; FILING DATE: 12-JAN-1998
; APPLICATION NUMBER: US/08/721,684
; FILING DATE: 27-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/33487
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 521 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-407-715-2

Query Match 100.0%; Score 2650; DB 4; Length 521;
Best Local Similarity 100.0%; Pred. No. 5.8e-215;
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MEVLSEGSQVLDNRKLSLSEPEGTALMYHFSSELLDEFSQVNLGQLLSDPFLSEK 60
Db 1 MEVLSEGSQVLDNRKLSLSEPEGTALMYHFSSELLDEFSQVNLGQLLSDPFLSEK 60
61 SESMEVEPSPTAPLIAEHSYSLSEPRQSPETHAATSDSFNDEEVESEKWLSTEF 120
Db 61 SESMEVEPSPTAPLIAEHSYSLSEPRQSPETHAATSDSFNDEEVESEKWLSTEF 120
Qy 121 PSATIKKEPITEEPPGLVPSVTLTITAIPTPFKEESPLDMNAGDSSCOTLIPIKILE 180
Db 121 PSATIKKEPITEEPPGLVPSVTLTITAIPTPFKEESPLDMNAGDSSCOTLIPIKILE 180
Qy 181 PHEVDQFLNFPKPEASVDQLHLPTPPSSHSDSGSLSPNPRLHPFSLQAHSPVRAMP 240
Db 181 PHEVDQFLNFPKPEASVDQLHLPTPPSSHSDSGSLSPNPRLHPFSLQAHSPVRAMP 240
Qy 241 RPSALSSTPLTAPHKLOGGSLVLTBEKRTLVAEGYPTKPLTKYSEKALKKIRR 300
Db 241 RPSALSSTPLTAPHKLOGGSLVLTBEKRTLVAEGYPTKPLTKYSEKALKKIRR 300
Qy 301 KIKNKISAESEKRRKKEKYMDSLEKKEVSCSTENLRKKVELENTRTLQLOKLOTL 360
Db 301 KIKNKISAESEKRRKKEKYMDSLEKKEVSCSTENLRKKVELENTRTLQLOKLOTL 360
Qy 361 VMGKVSRTCKLAGTGTGTCMLVMVLCFAVAFSGFPQGYGPSATKMLPSQHPLEPYT 420
Db 361 VMGKVSRTCKLAGTGTGTCMLVMVLCFAVAFSGFPQGYGPSATKMLPSQHPLEPYT 420
Qy 421 ASVVRNLLIYEEHAPLEESSSPASTGELGWDGSGSLLRASSGLEALPEVDLPFLIS 480
Db 421 ASVVRNLLIYEEHAPLEESSSPASTGELGWDGSGSLLRASSGLEALPEVDLPFLIS 480
481 NETSLEKSVLLEQOHLVSSKLEGNETLKVLELRRVNTF 521
Db 481 NETSLEKSVLLEQOHLVSSKLEGNETLKVLELRRVNTF 521

RESULT 4
US-09-247-155-113
Sequence 113, Application US/09247155A
Patent No. 6312922
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean-Baptiste
APPLICANT: Duclert, Aymeric
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: Complementary DNAs
FILE REFERENCE: GENSET.021A
CURRENT APPLICATION NUMBER: US/09/247,155A
CURRENT FILING DATE: 1999-02-09
EARLIER APPLICATION NUMBER: 60/074,121
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/081,563
EARLIER FILING DATE: 1998-04-13
EARLIER APPLICATION NUMBER: 60/096,116
EARLIER FILING DATE: 1998-08-10
EARLIER APPLICATION NUMBER: 60/099,273
EARLIER FILING DATE: 1998-10-04
NUMBER OF SEQ ID NOS: 182

SOFTWARE: Patent.pm
SEQ ID NO 113
LENGTH: 395
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -310..-1
US-09-247-155-113

Query Match 14.9%; Score 395; DB 4; Length 395;
Best Local Similarity 34.0%; Pred. No. 2.8e-25;
Matches 115; Conservative 49; Mismatches 108; Indels 66; Gaps 12;
Qy 135 PPGLVPSVTLTITAIPTPFKEESPLDMNAGDSSC--QTLP----KIKLEPHV--DQ 186
Db 33 PPPEVPVTRL-----QEQLGQWKGSGRGLQSEPEDFLKLFDNPNEYVYCE 82
Qy 187 FLNFPKPEASVDQLHLPTPPSSHSDS-----EGSLSPNPRLHPFSL 230
Db 83 ASPGSDSGISDSCH-PDPPAPRATSPMLYEVYEAALERMOGETGPNVGLISIQLD 141
Qy 231 QAHPVPRAMPGPSALSTSPLLTAPHKLOGSG-----PLVLTTEEEXRTLVA 276
Db 142 Q-WSPAFWVP-DSCMVSELFPDAHAHILPRAGTVAPVCTLLPCOTLFLTDEEKELCO 199
Qy 277 EGYPIPTKPLTKSEKALKKIRKIRKNIQAQSRKRRKKEYMDSLEKKEVSCSTENLEL 336
Db 200 EGVSLPSHLPTKAEERVLKVRKIRNKQAQDSRRRKEYIDGLSRVAACSAQNOEL 259
Qy 337 RKKVELENTRTLQLOKLOTLVMGKVSRTCKLAGTGTGTCMLVMVLCFAVAFSGFPQ 396
Db 260 QKKVQELERHNLISVAQLRQLOTL----IAQTSNKA-AQTSTCVLILLFLSLAILPLPSFS 314
Qy 397 GYGPVPSATKMLPSQHPLEPYTASVVRNLLIYEE 434
Db 315 PFQSRPEAG-----SEDIQPHGVTSRNLTHTKD 342

RESULT 5
US-07-684-965-6
Sequence 6, Application US/07684965
Patent No. 5919649
GENERAL INFORMATION:
APPLICANT: HABENER, JOEL F.
APPLICANT: HOEFLEER, JAMES P.
TITLE OF INVENTION: A-CAMP RESPONSIVE TRANSCRIPTIONAL
TITLE OF INVENTION: ENHANCER BINDING PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue, N.W., Suite 300
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/684,965
FILING DATE: 19910522
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0609.2470004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)833-7533
TELEFAX: (202)833-8716
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 327 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-684-965-6

Query Match 5.5%; Score 146.5; DB 2; Length 327;
Best Local Similarity 23.8%; Pred. No. 0.00018;
Matches 73; Conservative 52; Mismatches 105; Indels 77; Gaps 15;
QY 74 APLIAQHSYSLSEEPRTQSPFTHAATSDSFNDEEVESEK-WY-----LSTFPFSATIK 126
DB 80 SQVQTVQISTIAEHSQESVD--SVTDSQNRRLSRPSRYKRLNDLSDAFQVPI 137
QY 127 KSPITEQPGGLVPSVTLTITAISTPFKEESP--LDMNAGDSSCQTLIPKIKLEPHEV 184
DB 138 BEEKSEETS--APAIT-TVT-VPTPIYQTSQGVIATIQGG-----AIQLANNGT 184
QY 185 DQFLNFPKPEASVDQLHLPPTPS-----SHSDSEGLSPNPRHLHPFSLQAHSPVPA 238
DB 185 DGVQGLQ-----TLTWTNAAATPGTILQYACTTDOQQLVPTNQV---VVOAASGDVQT 237
QY 239 MPRGFSALSTPLLT-APHKLOGSGLVLTBEBKRTLVAEGYPIPTKLPKTKSEKALKK 297
DB 238 Y-----QIRTAFTSTIAPGVNMASPP-----ALPTQAEAPRKR 272
QY 298 IRRKINKISAQESRRKKKYMDSLEKVCSCSTENLELRKKVLENTNRTLQLOQLK 357
DB 273 EVRLMKNREARCRKKKYYVKCLE-----NRVAVLENQNTLIEBKAL 318
QY 358 QTLVMGK 364
DB 319 KOLYCHK 325

RESULT 6
US-09-215-098-5
Sequence 5, Application US/09215098
Patent No. 6194632
GENERAL INFORMATION:
APPLICANT: Leiden, Jeffery M
TITLE OF INVENTION: DILATED CARDIOMYOPATHY IN TRANSGENIC MICE EXPRESSING A
TITLE OF INVENTION: DOMINANT-NEGATIVE CREB TRANSCRIPTION FACTOR IN THE
FILE REFERENCE: 9189-4
CURRENT APPLICATION NUMBER: US/09/215,098
PRIOR FILING DATE: 1998-12-18
PRIOR APPLICATION NUMBER: 60/068,011
PRIOR FILING DATE: 1997-12-18
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 327
TYPE: PRT
ORGANISM: Homo sapiens
US-09-215-098-5

Query Match 5.5%; Score 146.5; DB 4; Length 327;
Best Local Similarity 22.6%; Pred. No. 0.00018;
Matches 85; Conservative 56; Mismatches 114; Indels 121; Gaps 18;
QY 50 QLLSDPFLSE-KGESMEVEPSP-----TSPAP-----L 76
DB 10 QQSGDAAVTAENQOMTVQAPQIATLAQVMPAAHATSGAPTTLVQLPNGQTVQVHGV 69
QY 77 IQAHSYSLSEEPRTQ-----SPFTHAATSDSFNDEEVESEK-WY-----LS 117
DB 70 IQAQP-SVQSPQVQTVQISTIAESDSESVSDTSQKRRILRRRYSYKILNDLS 128
QY 118 TEPFSATIKKPEITTEOPGGLVPSVTLTITAISTPFKEESP--LDMNAGDSSCQTLIP 175
DB 129 SDAPGVRIEESSEETS--APAIT-TVT-VPTPIYQTSQGVIATIQGG----- 175

QY 176 KIKLEPHEVDQFLNFPKPEASVDQLHLPPTPS-----SHSDSEGLSPNPRHLHPFSL 229
DB 176 AIQLANNGTGVQGLQ-----TLTWTNAAATPGTILQYACTTDOQQLVPTNQV---VV 228
QY 230 QAHSPPVRAMPGRPSALSTPLLT-APHKLOGSGLVLTBEBKRTLVAEGYPIPTKLP 288
DB 229 QAASGDVQTY-----QIRTAFTSTIAPGVNMASPP-----ALPTQ 263
QY 289 KSEKALKIKIRKINKISAQESRRKKKYMDSLEKVCSCSTENLELRKKVLENTNRTL 348
DB 264 PAEEAARREVRLMKNREARCRKKKYYVKCLE-----NRVAVLENQNTL 309
QY 349 TILQLOQLKQTLVMGK 364
DB 310 TLIEBKALKOLYCHK 325

RESULT 7
US-09-252-658-6
Sequence 6, Application US/09252658
Patent No. 6251667
GENERAL INFORMATION:
APPLICANT: HABENER, JOEL P.
APPLICANT: HOFFLER, JAMES P.
TITLE OF INVENTION: A CAMP-RESPONSIVE TRANSCRIPTIONAL
TITLE OF INVENTION: ENHANCER BINDING PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, STE. 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/252,658
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/684,965
FILING DATE: 22-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US89/05234
FILING DATE: 20-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/272,980
FILING DATE: 18-NOV-1988
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0609.1990002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 327 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-252-658-6

Query Match 5.5%; Score 146.5; DB 4; Length 327;
Best Local Similarity 23.8%; Pred. No. 0.00018;
Matches 73; Conservative 52; Mismatches 105; Indels 77; Gaps 15;
QY 74 APLIAQHSYSLSEEPRTQSPFTHAATSDSFNDEEVESEK-WY-----LSTFPFSATIK 126

Db 80 SPQVQTQVIAIESEHSESVVD--SVTDSQNRRLRRYSRKILNDLSDSDGCVPR1 137
Qy 127 KEPITEOPPGVLPSVTLTITAIPTPFKEBSP--LDMNAGDSSCQTLIPKIKLEPHEV 184
Db 138 EEKSEETS--APAIT-TVT-VPTPIYQTSQGVIAITQGG-----AIQLANNGT 184
Qy 185 DQFLNFSPEASVDOLHLPPTPS-----SHSSDSEGLSPNPLHPSLSQAHSPVRA 238
Db 185 DGVOGLQ---TLTWTNAAATOPGTTILOAYQTTDGOQILVPTNQV---VVOAASGDVQT 237
Qy 239 MPRGPSALSTPLLT-APHKLGSGPLVLTBEEKRTLVAEGVPIPTKLPLTKSEKALKK 297
Db 238 Y-----QIRTAFTSIAGVWASP-----ALPTQPAEEAPKR 272
Qy 298 IRRKIKISAOESRRKKKYNDLSLEKKVESCSTENLELRKKVEYLENTRTLLOQL 357
Db 273 EVRLMKNEARECRKKKYVKLE-----NRVALENQNTLIELKAL 318
358 QTLVNGK 364
319 KOLYCHK 325

RESULT 8
US-08-094-533B-10
; Sequence 10, Application US/08094533B
; Patent No. 5534426
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: MASAHICO, HIBI
; APPLICANT: ANNING, LIN
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/094,533B
; FILING DATE: 19-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/006001 (PD2848)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-094-533B-10

Query Match 5.5%; Score 146.5; DB 1; Length 331;
Best Local Similarity 20.3%; Pred. No. 0.00018;
Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;
Qy 68 PSPTSP-----APLIOAHSYSLSEEPRTQSPFTHAATSDSFNDEEV-----ESKW 114
Db 20 PSESGPYGYNPKILKQSMTLNADPVGSLPRLAKNSDLLTSPDVGLLKLASPELERL 79

Qy 115 YL-----STEPSATIKKEPITEOPPGVLPSVTLTITAIPTPFKEBSPDMDNAGG 166
Db 80 IQSSNGHITTTPTQFLCPKNVTDEQ-EGFAEGFVRAELHS----- 123
Qy 167 DSSCQTLIPKIKLEPHEVDQFLNFSPEASV-----DQLHLPTPPSSHSSDSEG 216
Db 124 ---QNTLPSVTSAAQPVNGAGWAPAVASVAGSGSGGFSASLHSEPPVYANLSNPNP 179
Qy 217 SLSNPRHLHPSLSQAHSPVPRGFSALSTPLLTAPHKLQSGSP-----LVLTEBEK 271
Db 180 ALSGGGAPSYGAAGL-----AFPAQPOQQO---QPPHLPQPMVQHPRLQALKEEP 229
Qy 272 RTLVAEGVPIPTKLPLTKSEKALKTKRKIKNISAOESRRKKKYNDLSLEKKVESCST 331
Db 230 QTVPEMGETPPLPIDMESQERIKAEKRMNRNIAASKCRKRLERLARLEEKVKTLKA 289
Qy 332 ENLELRKKVEYLENTRTLLOQLQKLOTLVNGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANWLREOVAQLKQKVNHNHNSGQQLMLTQ 326

RESULT 9
US-08-276-860A-10
; Sequence 10, Application US/08276860A
; Patent No. 5593884
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: MASAHICO, HIBI
; APPLICANT: ANNING, LIN
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,860A
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/017001 (PD3701)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-276-860A-10

Query Match 5.5%; Score 146.5; DB 1; Length 331;
Best Local Similarity 20.3%; Pred. No. 0.00018;
Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;
Qy 68 PSPTSP-----APLIOAHSYSLSEEPRTQSPFTHAATSDSFNDEEV-----ESKW 114
Db 20 PSESGPYGYNPKILKQSMTLNADPVGSLPRLAKNSDLLTSPDVGLLKLASPELERL 79
Qy 115 YL-----STEPSATIKKEPITEOPPGVLPSVTLTITAIPTPFKEBSPDMDNAGG 166
Db 80 IQSSNGHITTTPTQFLCPKNVTDEQ-EGFAEGFVRAELHS----- 123

Db 124 ----QNTLPSVTSAAQPVNGAGMVAPAVASVAGSGGGSFASLSHSBPPVYANLSNFPG 179

Qy 217 SLSPNRLHPFSLQASHSPVRAMPRGSPALSTSPLLTAPHKLGSGP-----LVLTBEK 271

Db 180 ALSGGGAPSAGAGL-----APPAQPOQQO-----QPPHLPQOMPVOHPRLOALKEEP 229

Qy 272 RTLVAEGVPIPTKPLTKSEBALKIRKIKNKISAOESRRKKKKEYMDSLEKKVESCST 331

Db 230 QTVEMPGETPLSPIDMESQERIKAEKRMNRNIAASKCRKRLERLARLEEKVKTLKA 289

Qy 332 ENLELRKKVELENTNRTLQLOKLOKLOTLVMKVSRTCKLAGTQ 375

Db 290 QNSE-----LASTANMLREQVAQLKQKVMNHVNSGQMLMTQ 326

RESULT 11

US-08-799-913-10

Sequence 10, Application US/08799913

Patent No. 5804399

GENERAL INFORMATION:

APPLICANT: KARIN, MICHAEL

APPLICANT: HIBI, MASAHIKO

APPLICANT: LIN, ANNING

TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: FISH & RICHARDSON P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: California

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/799,913

FILING DATE: 13-FEB-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/444,393

FILING DATE: 19-MAY-1995

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Ph.D., Lisa A.,

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07257/017002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 678-5070

TELEFAX: (619) 678-5099

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 331 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-799-913-10

Query Match 5.5%; Score 146.5; DB 1; Length 331;

Best Local Similarity 20.3%; Pred. No. 0.00018;

Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;

Qy 167 DSSCQTLIPKIKLEPHEVDQFLNFSPEASV-----DQLHLPPTPSSSHSDSEG 216

Db 124 ----QNTLPSVTSAAQPVNGAGMVAPAVASVAGSGGGSFASLSHSBPPVYANLSNFPG 179

Qy 217 SLSPNRLHPFSLQASHSPVRAMPRGSPALSTSPLLTAPHKLGSGP-----LVLTBEK 271

Db 180 ALSGGGAPSAGAGL-----APPAQPOQQO-----QPPHLPQOMPVOHPRLOALKEEP 229

Qy 272 RTLVAEGVPIPTKPLTKSEBALKIRKIKNKISAOESRRKKKKEYMDSLEKKVESCST 331

Db 230 QTVEMPGETPLSPIDMESQERIKAEKRMNRNIAASKCRKRLERLARLEEKVKTLKA 289

Qy 332 ENLELRKKVELENTNRTLQLOKLOKLOTLVMKVSRTCKLAGTQ 375

Db 290 QNSE-----LASTANMLREQVAQLKQKVMNHVNSGQMLMTQ 326

RESULT 10

US-444-393-10

Sequence 10, Application US/08444393

Patent No. 5605808

GENERAL INFORMATION:

APPLICANT: KARIN, MICHAEL

APPLICANT: HIBI, MASAHIKO

APPLICANT: LIN, ANNING

TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: FISH & RICHARDSON P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: California

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/444,393

FILING DATE: 19-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Ph.D., Lisa A.,

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07257/017002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 678-5070

TELEFAX: (619) 678-5099

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 331 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-444-393-10

Query Match 5.5%; Score 146.5; DB 1; Length 331;

Best Local Similarity 20.3%; Pred. No. 0.00018;

Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;

Qy 68 PSPTSP-----APLIQAEHSYLSSEPTQSPFTHAATSDSFNDEEV-----ESEKW 114

Db 20 PSBGPYGVNPKIKQSMNTLNADPVGSLKPHLRAXNSDLLTSPDVGLLKLASPELERL 79

Qy 115 YL-----STFPPSATIKKEPITERQPPGLVPSVTLTITATSTPPEKEESPLDMNAGG 166

Db 80 IQSSNGHIHTTPTQFLCPKNVTDEQ-EGFAGFVRALAEHLS-----123

Qy 167 DSSCQTLIPKIKLEPHEVDQFLNFSPEASV-----DQLHLPPTPSSSHSDSEG 216

Db 124 ---QNTLPSTVTSAAQPVNGAGWAPAVASVAGSGSGFSASLHSEPPVYANLSNFPNG 179
Qy 217 SLSPNRLHPFSLQAHSPVRAMPGRPSALSTPLTAPHKLGQSGP-----LVITEBEK 271
Db 180 ALSGGGAPSYGAAGL-----AFPAQPOQQ-----QPPHLLPQOMPVOHPRLQALKEEP 229
Qy 272 RTLVAEGYPIPTKLPLTKSEKALKIRKIKNKISAOESRRKKKEYMDSLEKKVESCST 331
Db 230 QTVPEMPGETPLSPIDMESQERIKAEKRWNRNRAASKCRKKLERTARLEBEKVTLKA 289
Qy 332 ENLELRKKVEVLNTRTLLOQLKQTLVMGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANMLREQVAQLKQVMNVHNSGCOLMLTQ 326

RESULT 12

US-08-711-893-10
; Sequence 10, Application US/08711893
; Patent No. 5837244
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: MASAHICO, HIBI
; APPLICANT: ANNING, LIN
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/711.893
; FILING DATE: 12-SEP-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276.860
; FILING DATE: 18-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/017001 (PD3701)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-711-893-10

Query Match 5.5%; Score 146.5; DB 2; Length 331;
Best Local Similarity 20.3%; Pred. No. 0.00018;
Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;

Qy 68 PSPTSP-----APLIQAHSYSLSEEPRTQSPFFTHAATSDFSNDVEV-----ESEKW 114
Db 20 PSESGPYGYNPKILKQSMTLNLDAPVGLKPLRAKNSDLTSPDVGLLKLASFELERL 79
Qy 115 YL-----STEPSATIKKEPITEEOPGLVPSVTITITAITPPEKESPLDMNAGG 166
Db 80 IIQSSNGHITTTPTPTQPLCPKNVTDEQ-EGFAEGFVRALAEHLS----- 123
Qy 167 DSSCQTLPKILKLPHEVDQFLNFSPEASV-----DQLHLPTPPSSHSDSEG 216

Db 124 ---QNTLPSTVTSAAQPVNGAGWAPAVASVAGSGSGFSASLHSEPPVYANLSNFPNG 179
Qy 217 SLSPNRLHPFSLQAHSPVRAMPGRPSALSTPLTAPHKLGQSGP-----LVITEBEK 271
Db 180 ALSGGGAPSYGAAGL-----AFPAQPOQQ-----QPPHLLPQOMPVOHPRLQALKEEP 229
Qy 272 RTLVAEGYPIPTKLPLTKSEKALKIRKIKNKISAOESRRKKKEYMDSLEKKVESCST 331
Db 230 QTVPEMPGETPLSPIDMESQERIKAEKRWNRNRAASKCRKKLERTARLEBEKVTLKA 289
Qy 332 ENLELRKKVEVLNTRTLLOQLKQTLVMGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANMLREQVAQLKQVMNVHNSGCOLMLTQ 326

RESULT 13

US-09-150-200-10
; Sequence 10, Application US/09150200
; Patent No. 5994513
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: HIBI, MASAHICO
; APPLICANT: LIN, ANNING
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/150.200
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/444.393
; FILING DATE: 19-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/017002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-150-200-10

Query Match 5.5%; Score 146.5; DB 2; Length 331;
Best Local Similarity 20.3%; Pred. No. 0.00018;
Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;

Qy 68 PSPTSP-----APLIQAHSYSLSEEPRTQSPFFTHAATSDFSNDVEV-----ESEKW 114
Db 20 PSESGPYGYNPKILKQSMTLNLDAPVGLKPLRAKNSDLTSPDVGLLKLASFELERL 79
Qy 115 YL-----STEPSATIKKEPITEEOPGLVPSVTITITAITPPEKESPLDMNAGG 166
Db 80 IIQSSNGHITTTPTPTQPLCPKNVTDEQ-EGFAEGFVRALAEHLS----- 123
Qy 167 DSSCQTLPKILKLPHEVDQFLNFSPEASV-----DQLHLPTPPSSHSDSEG 216

Db 124 ----QNTLPSTVSAAPVNGAGMVA PAVASVAGSGGGSFASLSHSEPPVYANLSNPNP 179

Qy 217 SLSNPRLHPFSLQAHSPVRAMPRGPSALSTSLTAPHKLQSGSP-----LVLTEEEK 271

Db 180 ALSSGGGAPSAGAGL-----APPAQPOQQO-----QPPHLPQOMPVOHPRLQALKEEP 229

Qy 272 RTLVAEGPIPTKPLTKSEBKALKIRKIKNKISAQESRRKKKEYMDSLEKKVESCST 331

Db 230 QTVPEMGETPPLSPIDMESQERIKAEKRMNRNIAASKCKRKLRIARLEEKVKTLKA 289

Qy 332 ENLELRKKVELVNTNRTLQLOKLOLTVMGKVSRTCKLAGTQ 375

Db 290 QNSE-----LASTANMLREQVAQLKOKVMNVHNSGCOLMLTQ 326

RESULT 15

US-09-452-370-10

; Sequence 10, Application US/09452370

; Patent No. 6193965

; GENERAL INFORMATION:

; APPLICANT: KARIN, MICHAEL

; APPLICANT: HIBI, MASAHIKO

; APPLICANT: LIN, ANNING

; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FISH & RICHARDSON P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: California

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09452,370

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/150,200

; FILING DATE:

; APPLICATION NUMBER: US 08/444,393

; FILING DATE: 19-MAY-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Haile, Ph.D., Lisa A.,

; REGISTRATION NUMBER: 38,347

; REFERENCE/DOCKET NUMBER: 07257/017002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 678-5070

; TELEFAX: (619) 678-5099

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 331 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-452-370-10

Query Match 5.5%; Score 146.5; DB 4; Length 331;

Best Local Similarity 20.3%; Pred. No. 0.00018;

Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;

Qy 68 PSPTSP-----APLQAEHSYSSEEPRTQSPFTHAATSDSFNDEEV-----ESEKW 114

Db 20 PSSEGPYGNPKILKOSMTNLADPVGSLKPHLRKNSDLTSPDVGLLKLASPELERL 79

Qy 115 YL-----STEPSATIKKEPITEEQPGLVPSVTLTITATSTPPEKESPLDMNAGG 166

Db 80 IQSSNGHITTTPTTQFLCPKNVTDEQ-EGFABGFVRLAELHS----- 123

Qy 167 DSSCQTLIPKLEPHEVDQFLNFSPEASV-----DQLHLPTTPSSHSSDSEG 216

Db 124 ----QNTLPSTVSAAPVNGAGMVA PAVASVAGSGGGSFASLSHSEPPVYANLSNPNP 179

Qy 217 SLSNPRLHPFSLQAHSPVRAMPRGPSALSTSLTAPHKLQSGSP-----LVLTEEEK 271

Db 180 ALSSGGGAPSAGAGL-----APPAQPOQQO-----QPPHLPQOMPVOHPRLQALKEEP 229

Qy 272 RTLVAEGPIPTKPLTKSEBKALKIRKIKNKISAQESRRKKKEYMDSLEKKVESCST 331

Db 230 QTVPEMGETPPLSPIDMESQERIKAEKRMNRNIAASKCKRKLRIARLEEKVKTLKA 289

Qy 332 ENLELRKKVELVNTNRTLQLOKLOLTVMGKVSRTCKLAGTQ 375

Db 290 QNSE-----LASTANMLREQVAQLKOKVMNVHNSGCOLMLTQ 326

RESULT 14

US-09-150-201-10

; Sequence 10, Application US/09150201

; Patent No. 6001584

; GENERAL INFORMATION:

; APPLICANT: KARIN, MICHAEL

; APPLICANT: HIBI, MASAHIKO

; APPLICANT: LIN, ANNING

; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FISH & RICHARDSON P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: California

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09150,201

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/444,393

; FILING DATE: 19-MAY-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Haile, Ph.D., Lisa A.,

; REGISTRATION NUMBER: 38,347

; REFERENCE/DOCKET NUMBER: 07257/017002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 678-5070

; TELEFAX: (619) 678-5099

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 331 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-150-201-10

Query Match 5.5%; Score 146.5; DB 3; Length 331;

Best Local Similarity 20.3%; Pred. No. 0.00018;

Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;

Qy 68 PSPTSP-----APLQAEHSYSSEEPRTQSPFTHAATSDSFNDEEV-----ESEKW 114

Db 20 PSSEGPYGNPKILKOSMTNLADPVGSLKPHLRKNSDLTSPDVGLLKLASPELERL 79

Qy 115 YL-----STEPSATIKKEPITEEQPGLVPSVTLTITATSTPPEKESPLDMNAGG 166

Db 80 IQSSNGHITTTPTTQFLCPKNVTDEQ-EGFABGFVRLAELHS----- 123

Qy 167 DSSCQTLIPKLEPHEVDQFLNFSPEASV-----DQLHLPTTPSSHSSDSEG 216

QY 167 SSSCOTLIPKIKLEPHEVDQFLNFSPEASV-----DQLHLPTTPSSHSSDSEG 216
Db 124 ----QNTLPSVTISAAQPVNGAGMVAFAVASAGSGSGGFSASLHSEPPVYANUSNPNPG 179
QY 217 SLSNPRLHPFSLSOAHSPVRAMPGRGPSALSTPLLTAPHKLGSGP-----LVLTEEEK 271
Db 180 ALSSGGGAPSYGAAGL-----AFPAQPOQQ-----QPPHLPQOMPVOHPRLQALKEEP 229
QY 272 RTLVAEGYPIPTKLPLTKSEBKALKIRKIKNKISAOESRRKKKKEYMDSLEKKVESCST 331
Db 230 QTVPEMPGETPPLSPIDMESQERIKAEKRMNRITAAKCRKRKLERTARLEEKVKTILKA 289
QY 332 ENLELRKKVELENTNRTLLOLOKLOTLVMGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANMLREQVAOLKOKVNVHNSGCCQLMLTQ 326

Search completed: March 5, 2003, 06:21:35
Job time : 64 secs

THIS PAGE BLANK (USPTO)

This Page is inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images
problems checked, please do not report the
problems to the IFW Image Problem Mailbox**

THIS PAGE BLANK (USPTO)